

GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Cheever, Martin A.  
APPLICANT: Hand-Zimmerman, Susan  
TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL  
FILE REFERENCE: 210121.483C3  
CURRENT APPLICATION NUMBER: US/10/313,644  
CURRENT FILING DATE: 2002-12-04  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-313-644-3

Query Match 100.0%; Score 54; DB 12; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00068;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EEVLVPOQGF 10  
DB 1 EEVLVPOQGF 10

RESULT 3  
US-09-854-356-5  
Sequence 5, Application US/09854356  
Patent No. US20020177567A1  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Gheysen, Dirk  
APPLICANT: Corixa Corporation  
APPLICANT: SmithKline Beecham Biologicals S. A.  
TITLE OF INVENTION: HER-2/neu Fusion Proteins  
FILE REFERENCE: 014058-009810PC  
CURRENT APPLICATION NUMBER: US/09/854,356  
CURRENT FILING DATE: 2001-05-09  
PRIOR APPLICATION NUMBER: US 09/493,480  
PRIOR FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: US 60/117,976  
PRIOR FILING DATE: 1999-01-25  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 59  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: fragment of the phosphorylation domain, preferred  
OTHER INFORMATION: portion (delta PD) of human HER-2/neu  
US-09-854-356-5

Query Match 100.0%; Score 54; DB 10; Length 59;  
Best Local Similarity 100.0%; Pred. No. 0.0046;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EEVLVPOQGF 10  
DB 32 EEVLVPOQGF 41

RESULT 4  
US-09-854-356-4  
Sequence 4, Application US/09854356  
Patent No. US20020177567A1  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Gheysen, Dirk  
APPLICANT: Corixa Corporation  
APPLICANT: SmithKline Beecham Biologicals S. A.  
TITLE OF INVENTION: HER-2/neu Fusion Proteins

FILE REFERENCE: 014058-009810PC  
CURRENT APPLICATION NUMBER: US/09/854,356  
CURRENT FILING DATE: 2001-05-09  
PRIOR APPLICATION NUMBER: US 09/493,480  
PRIOR FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: US 60/117,976  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 266  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu  
US-09-854-356-4

Query Match 100.0%; Score 54; DB 10; Length 266;  
Best Local Similarity 100.0%; Pred. No. 0.023;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EEVLVPOQGF 10  
DB 32 EEVLVPOQGF 41

RESULT 5  
US-10-102-806-583  
Sequence 583, Application US/10102806  
Publication No. US20030054421A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA103P1C1  
CURRENT APPLICATION NUMBER: US/10/102,806  
CURRENT FILING DATE: 2002-03-22  
PRIOR APPLICATION NUMBER: 09/925,298  
PRIOR FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05881  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 846  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 583  
LENGTH: 293  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (52)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (53)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (150)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (171)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (207)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (254)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-102-806-583

Query Match 100.0%; Score 54; DB 15; Length 293;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2004, 13:26:19 ; Search time 11 Seconds  
(without alignments)  
42.752 Million cell updates/sec

Title: US-09-930-125-3

Perfect score: 54

Sequence: 1 EBYLVPQGR 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID            | Description         |
|------------|-------|-------------|--------|---------------|---------------------|
| 1          | 54    | 100.0       | 245    | 1 ERB2_MOUSE  | P70424 mus musculu  |
| 2          | 54    | 100.0       | 1254   | 1 ERB2_MOUSE  | O60553 mesocricetu  |
| 3          | 54    | 100.0       | 1255   | 1 ERB2_HUMAN  | P04632 homo sapien  |
| 4          | 54    | 100.0       | 1257   | 1 ERB2_RAT    | P06494 rattus norv  |
| 5          | 50    | 92.6        | 1210   | 1 EGFR_HUMAN  | P00533 homo sapien  |
| 6          | 50    | 92.6        | 1210   | 1 EGFR_MOUSE  | Q01279 mus musculu  |
| 7          | 46    | 85.2        | 548    | 1 ERBB_AVIEU  | P11273 avian eryth  |
| 8          | 46    | 85.2        | 604    | 1 ERBB_AVIER  | P00535 avian eryth  |
| 9          | 46    | 85.2        | 634    | 1 ERBB_AVY    | P00533 avian leuko  |
| 10         | 38    | 70.4        | 232    | 1 YH73_SYNY3  | P73623 synechocyst  |
| 11         | 37    | 68.5        | 95     | 1 ERB4_MOUSE  | P73623 synechocyst  |
| 12         | 37    | 68.5        | 687    | 1 TGM2_BOVIN  | P51176 bos taurus   |
| 13         | 37    | 68.5        | 687    | 1 TGM2_HUMAN  | P11980 homo sapien  |
| 14         | 37    | 68.5        | 689    | 1 TGM2_CANCU  | P08587 cavia crite  |
| 15         | 37    | 68.5        | 1308   | 1 ERBB_HUMAN  | Q15303 homo sapien  |
| 16         | 37    | 68.5        | 1308   | 1 ERBB_RAT    | O62956 rattus norv  |
| 17         | 36    | 66.7        | 541    | 1 FTCD_HUMAN  | O62956 h. formimido |
| 18         | 35    | 64.8        | 391    | 1 TGM2_MOUSE  | P39101 mus musculu  |
| 19         | 35    | 64.8        | 666    | 1 TGM2_MOUSE  | P11981 mus musculu  |
| 20         | 35    | 64.8        | 917    | 1 NIA2_ARA/TH | P11035 arabidopsis  |
| 21         | 35    | 64.8        | 1242   | 1 MSH6_YEAT   | Q03844 saccharomyc  |
| 22         | 35    | 64.8        | 2269   | 1 WDR9_HUMAN  | Q03844 homo sapien  |
| 23         | 34    | 63.0        | 262    | 1 YES9_HELIPJ | O62956 h. formimido |
| 24         | 34    | 63.0        | 364    | 1 MURG_THETN  | O62956 thermococ    |
| 25         | 34    | 63.0        | 365    | 1 MURG_SHEVI  | O62956 thermococ    |
| 26         | 34    | 63.0        | 474    | 1 SYE_BUCAP   | O62956 thermococ    |
| 27         | 34    | 63.0        | 541    | 1 FTCD_PIG    | O62956 thermococ    |
| 28         | 34    | 63.0        | 598    | 1 NUSM_PETMA  | O62956 thermococ    |
| 29         | 34    | 63.0        | 988    | 1 TNP2_ECOLI  | P06694 escherichia  |
| 30         | 34    | 63.0        | 997    | 1 Y414_MYCPN  | P75183 mycoplasma   |
| 31         | 34    | 63.0        | 1097   | 1 TOLL_DROME  | P08953 drosophila   |
| 32         | 34    | 63.0        | 2290   | 1 POLG_BMCV   | P03304 encephalomy  |
| 33         | 34    | 63.0        | 2292   | 1 POLG_BMCV   | P17593 encephalomy  |

## ALIGNMENTS

|    |    |      |      |   |             |                    |
|----|----|------|------|---|-------------|--------------------|
| 34 | 34 | 63.0 | 2292 | 1 | POLG_BMCV   | P17594 encephalomy |
| 35 | 33 | 61.1 | 262  | 1 | YES9_HELIPJ | P55986 helicobacte |
| 36 | 33 | 61.1 | 370  | 1 | MYOM_APLCA  | P15513 a myomodul  |
| 37 | 33 | 61.1 | 390  | 1 | NCFI_MOUSE  | O09014 mus musculu |
| 38 | 33 | 61.1 | 392  | 1 | YAVI_SCHPO  | O10209 schizosacch |
| 39 | 33 | 61.1 | 439  | 1 | LIP1_DROME  | O46107 drosophila  |
| 40 | 33 | 61.1 | 459  | 1 | T7L2_MOUSE  | O97480 mus musculu |
| 41 | 33 | 61.1 | 541  | 1 | FTCD_MOUSE  | O91X40 m formimido |
| 42 | 33 | 61.1 | 541  | 1 | FTCD_RAT    | O88618 r formimido |
| 43 | 33 | 61.1 | 584  | 1 | T7L1_MOUSE  | O92131 mus musculu |
| 44 | 33 | 61.1 | 588  | 1 | T7L1_HUMAN  | O91954 homo sapien |
| 45 | 33 | 61.1 | 619  | 1 | T7L2_HUMAN  | O91954 homo sapien |

RESULT 1

ERB2\_MOUSE

ID ERB2\_MOUSE STANDARD; PRT; 245 AA.

AC P70424; Q61525;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Receptor (protein-tyrosine kinase erbB-2 (BC 2.7.1.112) (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (fragments).

GN ERBB2 OR NEU.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxId=10090;

RN [1]

RP SEQUENCE OF 1-149 FROM N.A.

RC STRAIN=CD-1; TISSUE=Uterus;

RX MEDLINE=97200814; PubMed=9048643;

RA Lip J., Day S.K., Day S.K.;

RT "Differential expression of the erbB2 gene in the perimplantation mouse uterus: potential mediator of signaling by epidermal growth factor-like growth factors.";

RT Endocrinology 138:1328-1337(1997).

EN [2]

RP SEQUENCE OF 150-245 FROM N.A.

RX MEDLINE=96069911; PubMed=7589796;

RA Mosecoso L.M., Chu G.C., Gautam M., Noakes P.G., Merlie J.P., Sames J.R.;

RT "Synapse-associated expression of an acetylcholine receptor-inducing protein, ARIA/hergulin, and its putative receptors, ErbB2 and ErbB3, in developing mammalian muscle.";

RT Dev. Biol. 172:158-169(1995).

RN [3]

RP INTERACTION WITH PRKCAP.

RX MEDLINE=2126773; PubMed=11278603;

RA Jaulin-Bastard F., Saito H., Le Bivic A., Ollendorff V., Marchetto S., Brindum D., Borg J.-P.;

RT "The ERBB2/HER2 receptor differentially interacts with ERBIN and PICK1 PSD-95/Dlg-20-1 domain proteins.";

RT J. Biol. Chem. 276:15256-15263(2001).

CC -1- FUNCTION: Essential component of a neuroligin-receptor complex, although neuroligins do not interact with it alone. Gp30 is a potential ligand for this receptor. Not activated by EGF, TGF-alpha and amphiregulin.

CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.

CC -1- SUBUNIT: Heterodimer with each of the other ERBB receptors (potential). Interacts with PRKCAP.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: Expressed predominantly in uterine epithelial cells. In the muscle, expression localizes to the synaptic sites of muscle fibers.

CC -1- DEVELOPMENTAL STAGE: On days 1-4 of pregnancy, ERBB2 is detected primarily in epithelial cells, the day 1 uterus showing the highest accumulation. On day 5, the epithelium and the decidualizing stromal cells around the implanting blastocyst

exhibit accumulation of this receptor. On days 6-8, the expression  
persists in the epithelium at both the implantation and  
interimplantation sites in addition to modest levels in the  
secondary decidua zone. On days 7 and 8, accumulation is also  
prominent in the trophoblastic giant cells.  
-1- PFM: Ligand-binding increases phosphorylation on tyrosine  
residues (by similarity).  
-1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; U71126; AAB17380.1; -;  
EMBL; L47239; AAB3532.1; -;  
HSSP; P11362; 1FGK.  
DR GO; GO:0007507; P:heart development; IMP.  
DR GO; GO:0007422; P:peripheral nervous system development; IMP.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR PRINTS; PR00109; TYRKINASE.  
DR PROSITE; PS000001; Prot\_kinase; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; PARTIAL.  
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
KW Transmembrane; Glycoprotein; Multigene family; Receptor; Transferase;  
KW Tyrosine-protein kinase; ATP-binding; Phosphorylation.  
FT NON\_TER 1 1  
FT DOMAIN <1 >149 PROTEIN KINASE.  
FT ACT\_SITE 61 61 BY SIMILARITY.  
FT NON\_CONS 149 150  
FT NON\_TER 245 245  
SQ SEQUENCE 245 AA; 26927 MW; 0F763F0363DFEFLC CRC64;  
  
Query Match 100.0%; Score 54; DB 1; Length 245;  
Best Local Similarity 100.0%; Pred. NO. 0.0021;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 EBYLVPOQGF 10  
DB 160 EBYLVPOQGF 169  
  
RESULT 2  
ERR2\_MESAU STANDARD; PRT; 1254 AA.  
ID ERR2\_MESAU  
AC 06053;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Receptor protein-tyrosine kinase erbB-2 precursor (BC 2.7.1.112)  
DE (p185erbB2) (neu proto-oncogene) (C-erbB-2).  
GN ERBB2 OR NEU.  
OS Mesocricetus auratus (Golden hamster).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Mesocricetus.  
OX NCB1\_TaxId=10036;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Neve;  
RX MEDLINE=94191007; PubMed=7908275;  
RA Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M.,  
RA Yamazaki Y., Ishikawa T.;  
RT Cloning and activation of the Syrian hamster neu proto-oncogene.";  
RU Gene 140;251-255(1994).  
CC -1- FUNCTION: Essential component of a neurogulin-receptor complex,  
although neurogulins do not interact with it alone. GP30 is a

potential ligand for this receptor. Not activated by EGF, TGF-  
alpha and amphiregulin (by similarity).  
-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
tyrosine phosphate.  
-1- SUBUNIT: Heterodimer with each of the other ERBB receptors  
(Potential). Interacts with PRKCAP (by similarity).  
-1- SUBCELLULAR LOCATION: Type I membrane protein.  
-1- PFM: Ligand-binding increases phosphorylation on tyrosine  
residues.  
-1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; D16295; BAA03801.1; -;  
PIR; I48161; I48161.  
DR HSSP; P11362; 1FGK.  
DR InterPro; IPR000494; EGFR\_L domain.  
DR InterPro; IPR006211; Furin-Like.  
DR InterPro; IPR006212; Furin\_repeat.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR004019; YLP motif.  
DR Pfam; PF00757; Furin-like; 1.  
DR Pfam; PF00069; kinase; 1.  
DR Pfam; PF01030; Recep\_L\_domain; 2.  
DR Pfam; PF02757; YLP; 2.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Prot\_kinase; 1.  
DR SMART; SM00261; FU; 4.  
DR SMART; SM00219; TYKFC; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
KW Proto-oncogene; Disease mutation.  
FT SIGNAL 1 21  
FT CHAIN 22 1254  
FT DOMAIN 22 652  
FT TRANSMEM 653 675  
FT DOMAIN 676 1254  
FT DOMAIN 158 368  
FT DOMAIN 472 644  
FT DOMAIN 720 987  
FT NP\_BIND 726 734  
FT BINDING 753 753  
FT ACT\_SITE 845 845  
FT DISULFID 195 204  
FT DISULFID 199 212  
FT DISULFID 236 244  
FT DISULFID 240 252  
FT DISULFID 255 264  
FT DISULFID 268 295  
FT DISULFID 299 311  
FT DISULFID 315 331  
FT DISULFID 334 338  
FT DISULFID 511 520  
FT DISULFID 515 528  
FT DISULFID 531 540  
FT DISULFID 544 560  
FT DISULFID 563 576  
FT DISULFID 567 584  
FT DISULFID 587 596  
FT DISULFID 600 623  
FT DISULFID 626 634  
FT DISULFID 630 642  
FT MOD\_RES 1139 1139  
PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

FT MOD RES 1247 1247 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 658 658 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 659 659 V -> E (IN ONCOGENIC NEU).  
 FT VARIANT 659 659 V -> R (IN ONCOGENIC NEU).  
 SO SEQUENCE 1254 AA; 138252 MW; 974C3791C21F2BBE1 CRC64;

Query Match 100.0%; Score 54; DB 1; Length 1254;  
 Best Local Similarity 100.0%; Pred. No. 0.011;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EBYLVPOQGF 10  
 DB 1021 EBYLVPOQGF 1030

RESULT 3  
 EBB2\_HUMAN STANDARD; PRT; 1255 AA.  
 ID EBB2\_HUMAN STANDARD; PRT; 1255 AA.  
 AC P04626;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last annotation update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)  
 DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell  
 DE surface receptor HER2) (MLN 19).  
 GN EBB2 OR HER2 OR NGL OR NEU.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID:9606;  
 RX MEDLINE:86118663; PubMed:3003577;  
 RA Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,  
 RA Salto T., Toyoshima K.;  
 RA "Similarity of protein encoded by the human c-erbB-2 gene to  
 RA epidermal growth factor receptor.";  
 RA Nature 319:230-234 (1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE:86070181; PubMed:2999974;  
 RA Cousens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,  
 RA McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,  
 RA Francke U., Levinson A., Ullrich A.;  
 RA "Tyrosine kinase receptor with extensive homology to EGF receptor  
 RA shares chromosomal location with neu oncogene.";  
 RA Science 230:1132-1139 (1985).  
 RN [3]  
 RP SEQUENCE OF 737-1031 FROM N.A.  
 RX MEDLINE:86016729; PubMed:299567;  
 RA Semba K., Kamata N., Toyoshima K., Yamamoto T.;  
 RA "A v-erbB-related protooncogene, c-erbB-2, is distinct from the  
 RA c-erbB-1/epidermal growth factor-receptor gene and is amplified in a  
 RA human salivary gland adenocarcinoma.";  
 RA Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501 (1985).  
 RN [4]  
 RP VARIANTS VAL-654 AND VAL-655  
 RX MEDLINE:93194196; PubMed:8095488;  
 RA Ehsani A., Low J., Wallace R.B., Wu A.M.;  
 RA "Characterization of a new allele of the human EBB2 gene by allele-  
 RA specific competition hybridization.";  
 RA Genomics 15:426-429 (1993).  
 CC -1- FUNCTION: Essential component of a neuroguilin-receptor complex,  
 CC although neuroguiline do not interact with it alone. G930 is a  
 CC potential ligand for this receptor. Not activated by EGF, TGF-  
 CC alpha and amphiregulin.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein

CC tyrosine phosphatase.  
 CC -1- SUBUNIT: Heterodimer with each of the other EBB2 receptors  
 CC (Potential). Interacts with PRKCAP (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- PTM: Ligand-binding increases phosphorylation on tyrosine  
 CC residues (By similarity).  
 CC -1- POLYMORPHISM: There are four alleles due to the variations in  
 CC positions 654 and 655. Allele B1 (654-Ile-Ile-655) has a frequency  
 CC of 0.782; allele B2 (654-Ile-Val-655) has a frequency of 0.206;  
 CC allele B3 (654-Val-Val-655) has a frequency of 0.012.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M1767; AAA35808.1; -;  
 CC EMBL; M1764; AAA35808.1; JOINED.  
 CC EMBL; M1763; AAA35808.1; JOINED.  
 CC EMBL; M1764; AAA35808.1; JOINED.  
 CC EMBL; M1765; AAA35808.1; JOINED.  
 CC EMBL; M1766; AAA35808.1; JOINED.  
 CC EMBL; M1730; AAA75493.1; -;  
 CC EMBL; M12036; AAA35978.1; -;  
 CC EMBL; X03363; CAA27060.1; -;  
 CC PIR; A24571; A24571.  
 CC PDB; 1OR1; 18-FEB-03.  
 CC DR PDB; 1OR1; 01-JAN-00.  
 CC DR GeneW; HGNC:3430; EBB2.  
 CC MIM; 164870; -;  
 CC GO; GO:0005012; F:Neu/ErbB-2 receptor activity; TAS.  
 CC GO; GO:0004716; F:receptor signaling protein tyrosine kinase . . . ; TAS.  
 CC GO; GO:0008283; P:cell proliferation; TAS.  
 CC GO; GO:0007048; P:progenesis; TAS.  
 CC GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.  
 CC GO; GO:0006468; P:protein amino acid phosphorylation; TAS.  
 CC InterPro; IPR000494; EGFR\_L\_domain.  
 CC InterPro; IPR006211; Furin-like.  
 CC InterPro; IPR000719; Prot Kinase.  
 CC InterPro; IPR001245; Tyr Kinase.  
 CC InterPro; IPR004019; YLP\_motif.  
 CC Pfam; PF00757; Furin-like; 1.  
 CC Pfam; PF00069; pkinase; 1.  
 CC Pfam; PF01030; Recep\_L\_domain; 2.  
 CC Pfam; PF02757; YLP; 2.  
 CC PRINTS; PRO0109; TYRKINASE.  
 CC ProDom; PD000001; Prot\_Kinase; 1.  
 CC SMART; SM00261; FU; 4.  
 CC DR SMART; SM00219; TYKC; 1.  
 CC DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 CC DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
 CC DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
 CC KX Transferrin; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 CC KM Polymorphism; 3d-structure.  
 CC FT SIGNAL 1 21  
 CC FT CHAIN 22 1255  
 CC FT DOMAIN 22 652  
 CC FT TRANSMEM 653 675  
 CC FT DOMAIN 676 1255  
 CC FT NP\_BIND 720 987  
 CC FT BINDING 726 734  
 CC FT ACT\_SITE 753 753  
 CC FT ACT\_SITE 845 845  
 CC FT DISULFID 195 204  
 CC FT DISULFID 199 212  
 CC FT DISULFID 220 227  
 CC POTENTIAL.  
 CC RECEPTOR PROTEIN-TYROSINE KINASE EBB2-2.  
 CC EXTRACELLULAR (POTENTIAL).  
 CC POTENTIAL.  
 CC CYTOPLASMIC (POTENTIAL).  
 CC PROTEIN KINASE.  
 CC ATP (BY SIMILARITY).  
 CC ATP (BY SIMILARITY).  
 CC BY SIMILARITY.  
 CC BY SIMILARITY.  
 CC BY SIMILARITY.  
 CC BY SIMILARITY.

```

FT DISULFID 224 235 BY SIMILARITY.
FT DISULFID 226 244 BY SIMILARITY.
FT DISULFID 240 252 BY SIMILARITY.
FT DISULFID 255 264 BY SIMILARITY.
FT DISULFID 268 295 BY SIMILARITY.
FT DISULFID 299 311 BY SIMILARITY.
FT DISULFID 315 331 BY SIMILARITY.
FT DISULFID 334 338 BY SIMILARITY.
FT DISULFID 511 520 BY SIMILARITY.
FT DISULFID 515 528 BY SIMILARITY.
FT DISULFID 531 540 BY SIMILARITY.
FT DISULFID 544 560 BY SIMILARITY.
FT DISULFID 563 576 BY SIMILARITY.
FT DISULFID 567 584 BY SIMILARITY.
FT DISULFID 587 596 BY SIMILARITY.
FT DISULFID 600 623 BY SIMILARITY.
FT DISULFID 626 634 BY SIMILARITY.
FT DISULFID 630 642 BY SIMILARITY.
FT MOD_RES 1139 1139 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1248 1248 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 654 654 I -> V (IN dbsnp:1801201).
FT VARIANT 655 655 I -> V (IN dbsnp:1801200).
FT VARIANT 655 655 /FtId=VAR_004078.
FT CONFLICT 1170 1170 P -> A (IN REF. 2).
FT SEQUENCE 1255 AA; 137909 MW; 39E9DFDA04DCF962 CRC64;

Query Match 100.0%; Score 54; DB 1; Length 1255;
Beer Local Similarity 100.0%; Pred No. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPQGGF 10
DB 1021 EBYLVPQGGF 1030

RESULT 4
ERRB2 RAT STANDARD; PRT; 1257 AA.
ID _ERRB2 RAT
AC P06454;
DT 01-JAN-1988 (Rel. 06, Created)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-2 precursor (BC 2.7.1.112)
DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Epidermal growth factor
DE receptor-related protein).
CN ERBB2 OR NEU.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Neuroblastoma;
RC MEDLINE=86118662; PubMed=3945311;
RA Bergmann C.I., Hung M.-C., Weinberg R.A.;
RT "The neu oncogene encodes an epidermal growth factor receptor-related
RT protein.";
RL Mature 319:226-230(1986).
RN [2]
RP SEQUENCE OF 852-905 FROM N.A.
RC TISSUE=Sciatic nerve;
RC MEDLINE=91222560; PubMed=2025425;
RA Lai C., Lemke G.;
RT "An extended family of protein-tyrosine kinase genes differentially
RT expressed in the vertebrate nervous system.";

```

```

RL Neuron 6:691-704(1991).
RN [3]
RP STRUCTURE BY NMR OF 650-668.
RX MEDLINE=92155181; PubMed=1346763;
RA Guillek W.J., Bottomley A.C., Loftis F.J., Doak D.G., Mulvey D.,
RA Newman R., Crumpton M.J., Sternberg M.J.B., Campbell I.D.;
RT "Three dimensional structure of the transmembrane region of the proto-
RT oncogenic and oncogenic forms of the neu protein.";
RL EMBO J. 11:43-48(1992).
CC -I- FUNCTION: Essential component of a neurogulin-receptor complex,
CC although neurogulins do not interact with it alone. GP30 is a
CC potential ligand for this receptor. Not activated by EGF, TGF-
CC alpha, and amphiregulin.
CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -I- SUBUNIT: Heterodimer with each of the other ERBB receptors. The
CC constitutively activated oncogenic variant forms a homodimer.
CC Interacts with PKCABP (By similarity).
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- PTM: Ligand-binding increases phosphorylation on tyrosine
CC residues (By similarity).
CC -I- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X03362; CA27059.1; ALT_INIT.
CC PDB: 1IIJ; 27-JUN-01.
CC PDB: 1N8Y; 18-FEB-03.
CC InterPro: IPR000494; EGFR_L domain.
CC InterPro: IPR006211; Furin-like.
CC InterPro: IPR006212; Furin repeat.
CC InterPro: IPR000719; Prot_Kinase.
CC InterPro: IPR001245; Tyr_kinase.
CC InterPro: IPR004019; YLP motif.
CC Pfam: PF00757; Furin-like; 1.
CC Pfam: PF00069; kinase; 1.
CC Pfam: PF01030; Recep_L_domain; 2.
CC Pfam: PF02757; YLP; 2.
CC PRINTS: PR00109; TYRKINASE.
CC PRODOM: PD000001; Prot_kinase; 1.
CC SMART: SM00261; FU; 4.
CC DR SMART; SM00219; Tyrc; 1.
CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
CC DR Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
CC KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
CC KW Proto-oncogene; Disease mutation; 3d-structure.
CC FT SIGNAL 1 21 POTENTIAL.
CC FT CHAIN 22 1257 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
CC FT DOMAIN 22 654 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 655 677 POTENTIAL.
CC FT DOMAIN 678 1257 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 159 369 CYS-RICH.
CC FT DOMAIN 473 646 CYS-RICH.
CC FT DOMAIN 722 969 PROTEIN KINASE.
CC FT NP BIND 728 736 ATP (BY SIMILARITY).
CC FT BINDING 755 755 ATP (BY SIMILARITY).
CC FT ACT SITE 847 847 BY SIMILARITY.
CC FT DISULFID 196 205 BY SIMILARITY.
CC FT DISULFID 200 213 BY SIMILARITY.
CC FT DISULFID 221 228 BY SIMILARITY.
CC FT DISULFID 225 236 BY SIMILARITY.
CC FT DISULFID 237 245 BY SIMILARITY.
CC FT DISULFID 241 253 BY SIMILARITY.
CC FT DISULFID 256 265 BY SIMILARITY.
CC FT DISULFID 269 296 BY SIMILARITY.

```

FT DISULFID 300 312 BY SIMILARITY.  
 FT DISULFID 316 332 BY SIMILARITY.  
 FT DISULFID 335 339 BY SIMILARITY.  
 FT DISULFID 513 522 BY SIMILARITY.  
 FT DISULFID 517 530 BY SIMILARITY.  
 FT DISULFID 533 542 BY SIMILARITY.  
 FT DISULFID 546 562 BY SIMILARITY.  
 FT DISULFID 565 578 BY SIMILARITY.  
 FT DISULFID 569 586 BY SIMILARITY.  
 FT DISULFID 589 598 BY SIMILARITY.  
 FT DISULFID 602 625 BY SIMILARITY.  
 FT DISULFID 628 636 BY SIMILARITY.  
 FT DISULFID 632 644 BY SIMILARITY.  
 FT MOD RES 1141 1141 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD RES 1250 1250 PHOSPHORYLATION (BY SIMILARITY).  
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 661 661 V -> E (IN ONCOGENIC NEU).  
 SO SEQUENCE 1257 AA; 138831 MW; 6129264583011402 CRC64;  
 Query Match 100.0%; Score 54; DB 1; Length 1257;  
 Best Local Similarity 100.0%; Pred. No. 0.011;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EBYLVPQGGP 10  
 Db 1023 EBYLVPQGGP 1032  
 RESULT 5  
 EGFR HUMAN STANDARD; PRT; 1210 AA.  
 ID EGFR\_HUMAN P00533; O00688; O00732; P06268; Q14225; Q92795; Q9B2S2; Q9GZX1;  
 AC P00533; O00688; O00732; P06268; Q14225; Q92795; Q9B2S2; Q9GZX1;  
 AC Q9H2C9; Q9H3C9; Q9H3C9; Q9H3C9; Q9H3C9; Q9H3C9;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Receptor  
 protein-tyrosine kinase ErbB-1).  
 GN EGFR OR ERBB1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RX MEDLINE=84219729; PubMed=6328312;  
 RA Ullrich A., Coussens L., Hayflick J.S., Dull T.J., Gray A., Tam A.W.,  
 Lee J., Yarden Y., Libermann T.A., Schlessinger J., Downward J.,  
 Mayer E.L.V., Whittle N., Waterfield M.D., Seeburg P.H.;  
 RA "Human epidermal growth factor receptor cDNA sequence and aberrant  
 expression of the amplified gene in A431 epidermoid carcinoma cells.";  
 RT Nature 309:418-425(1984).  
 RN [12]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Placenta;  
 RA MEDLINE=95382957; PubMed=7654368;  
 RA Ilekis J.V., Stark B.C., Scoccia B.;  
 RT "Possible role of variant RNA transcripts in the regulation of  
 epidermal growth factor receptor expression in human placenta.";  
 RL Mol. Reprod. Dev. 41:149-156(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Placenta;  
 RA MEDLINE=97078686; PubMed=8918811;  
 RA Reiter J.L., Mathie N.J.;  
 RT "A 1.8 kb alternative transcript from the human epidermal growth  
 factor receptor gene encodes a truncated form of the receptor.";  
 RL Nucleic Acids Res. 24:4050-4056(1996).

RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Placenta;  
 RA MEDLINE=97256547; PubMed=9103388;  
 RA Ilekis J.V., Garcil J., Niederberger C., Scoccia B.;  
 RT "Expression of a truncated epidermal growth factor receptor-like  
 protein (TEGFR) in ovarian cancer.";  
 RL Gynecol. Oncol. 65:36-41(1997).  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).  
 RC TISSUE=Placenta;  
 RA MEDLINE=21100872; PubMed=11161793;  
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,  
 Scheel Sinclair C., Pearse R.S., Green P.J., Yee D., Lampend A.L.,  
 Balasubramanian S., Crossley T.D., Magnuson T.R., James C.D.,  
 Mathie N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and  
 mouse alternative EGFR transcripts encoding truncated receptor  
 isoforms.";  
 RL Genomics 71:1-20(2001).  
 RN [6]  
 RP SEQUENCE OF 575-687 FROM N.A.  
 RA Reiter J.L., Threadgill D.W., Danielsen A.J., Scheel C.M.,  
 Lampend A.L., Balasubramanian S., Crossley T.O., Magnuson T.R.,  
 Mathie N.J.;  
 RT "Human and mouse alternative EGFR transcripts encoding only the  
 extracellular domain of the receptor.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE OF 713-924 FROM N.A.  
 RA MEDLINE=84196372; PubMed=6326261;  
 RA Lin C.R., Chen W.S., Kruliger W., Stolarsky L.S., Weber W.,  
 Evans R.M., Verma I.M., Gill G.N., Rosenfeld M.G.;  
 RT "Expression cloning of human EGF receptor complementary DNA: gene  
 amplification and three related messenger RNA products in A431  
 cells.";  
 RL Science 224:843-848(1984).  
 RN [8]  
 RP SEQUENCE OF 150-962 FROM N.A.  
 RA MEDLINE=84245835; PubMed=6330563;  
 RA Xu Y.H., Iehli S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P.,  
 Roe B.A., Merlino G.T., Pastan I.;  
 RT "Human epidermal growth factor receptor cDNA is homologous to a  
 variety of RNAs overproduced in A431 carcinoma cells.";  
 RL Nature 309:806-810(1984).  
 RN [9]  
 RP SEQUENCE OF 1028-1210 FROM N.A.  
 RA MEDLINE=85046483; PubMed=6093780;  
 RA Simmen F.A., Gope M.L., Schulz T.Z., Wright D.A., Carpenter G.,  
 O'Malley B.W.;  
 RT "Isolation of an evolutionarily conserved epidermal growth factor  
 receptor cDNA from human A431 carcinoma cells.";  
 RL Biochem. Biophys. Res. Commun. 124:125-132(1984).  
 RN [10]  
 RP SEQUENCE OF 1-29 FROM N.A.  
 RA MEDLINE=88217333; PubMed=3329716;  
 RA Haley J.D., Whittle N., Bennett P., Kinchington D., Ullrich A.,  
 Waterfield M.D.;  
 RT "The human EGF receptor gene: structure of the 110 kb locus and  
 identification of sequences regulating its transcription.";  
 RL Oncogene Res. 1:375-396(1987).  
 RN [11]  
 RP SEQUENCE OF 1-29 FROM N.A.  
 RA MEDLINE=91107677; PubMed=1988448;  
 RA Haley J.D., Waterfield M.D.;  
 RT "Contributory effects of de novo transcription and premature  
 transcript termination in the regulation of human epidermal growth  
 factor receptor proto-oncogene RNA synthesis.";  
 RL J. Biol. Chem. 266:1746-1753(1991).  
 RN [12]  
 RP SEQUENCE OF 1-29 FROM N.A.  
 RA MEDLINE=85270438; PubMed=2991899;  
 RA Iehli S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.;

RT "Characterization and sequence of the promoter region of the human  
RT epidermal growth factor receptor gene.";  
CC Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924 (1985).  
RN [13]  
RP SEQUENCE OF 540.  
RA Kohda D.;  
RL Submitted (SEP-1997) to the SWISS-PROT data bank.  
RN [14]  
RP RECEPTOR ACTIVITY.  
RX MEDLINE=84191554; PubMed=6325948;  
RA Mroczkowski B., Mosig G., Cohen S.;  
RT "ATP-stimulated interaction between epidermal growth factor receptor  
RT and supercoiled DNA.";  
RL Nature 309:270-273 (1984).  
RN [15]  
RP PHOSPHORYLATION.  
RX MEDLINE=89278137; PubMed=2543678;  
RA Margolis B.L., Lax I., Kris R., Dombalagian M., Honegger A.M.,  
RA Howk R., Givol D., Ullrich A., Schlessinger J.;  
RT "All autophosphorylation sites of epidermal growth factor (EGF)  
RT receptor and HER2/neu are located in their carboxyl-terminal tails.  
RT Identification of a novel site in EGF receptor.";  
RL J. Biol. Chem. 264:10667-10671 (1989).  
RN [16]  
RP CARBOHYDRATE-LINKAGE SITES ASN-128; ASN-175; ASN-413; ASN-444 AND  
RP ASN-528.  
RX MEDLINE=96398132; PubMed=8962717;  
RA Smilh K.D., Davies M.J., Bailey D., Renouf D.V., Hounsell E.F.;  
RT "Analysis of the glycosylation patterns of the extracellular domain of  
RT the epidermal growth factor receptor expressed in Chinese hamster  
RT ovary fibroblasts.";  
RL Growth Factors 13:121-132 (1996).  
RN [17]  
RP CARBOHYDRATE-LINKAGE SITES ASN-56; ASN-352; ASN-361; ASN-568 AND  
RP ASN-603.  
RX MEDLINE=20198209; PubMed=10731668;  
RA Sato C., Kim J.-H., Abe Y., Saito K., Yokoyama S., Kohda D.;  
RT "Characterization of the N-Oligosaccharides attached to the atypical  
RT Asn-X-Cys sequence of recombinant human epidermal growth factor  
RT receptor.";  
RL J. Biochem. 127:65-72 (2000).  
RN [18]  
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.  
RX MEDLINE=98225196; PubMed=9556602;  
RA Abe Y., Ooka M., Inagaki F., Lax I., Schlessinger J., Kohda D.;  
RT "Disulfide bond structure of human epidermal growth factor receptor.";  
RL J. Biol. Chem. 273:11150-11157 (1998).  
RN [19]  
RP REVIEW.  
RX MEDLINE=87297456; PubMed=3039909;  
RA Carpenter G.;  
RT "Receptors for epidermal growth factor and other polypeptide  
RT mitogens.";  
RL Annu. Rev. Biochem. 56:881-914 (1987).  
RN [20]  
RP INTERACTION WITH RPKL.  
RX PubMed=1116146;  
RA Habbib A.A., Chatterjee S., Park S.-K., Ratan R.R., Lefebvre S.,  
RA Vartanian T.;  
RT "The epidermal growth factor receptor engages receptor interacting  
RT protein and nuclear factor-kappa B (NF-kappa B)-inducing kinase to  
RT activate NF-kappa B. Identification of a novel receptor-tyrosine  
RT kinase signalosome.";  
RL J. Biol. Chem. 276:8865-8874 (2001).  
RN [21]  
RP FUNCTION: Receptor for EGF, but also for other members of the EGF  
CC family, as TGF-alpha, amphiregulin, betacellulin, heparin-binding  
CC EGF-like growth factor, gp30 and vaccinia virus growth factor. Is  
CC involved in the control of cell growth and differentiation.  
CC -1- FUNCTION: Isoform 2/truncated isoform may act as an antagonist.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -1- SUBUNIT: Binds RPKL.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Isoform 2 is

CC secreted.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=4;  
CC Name=1; Synonyms=pl170;  
CC Name=2; Synonyms=ep60; Truncated, TEGFR;  
CC Name=3; Synonyms=pl110;  
CC Name=4; Synonyms=VSP\_002887, VSP\_002890;  
CC Name=5; Synonyms=VSP\_002889, VSP\_002892;  
CC -1- TISSUE SPECIFICITY: Expressed in placenta. Isoform 2 is also  
CC expressed in ovarian cancers.  
CC -1- MISCELLANEOUS: Binding of EGF to the receptor leads to  
CC dimerization, internalization of the EGF-receptor complex,  
CC induction of the tyrosine kinase activity, stimulation of cell DNA  
CC  
Query Match 92.6%; Score 50; DB 1; Length 1210;  
Best Local Similarity 80.0%; Pred. No. 0.066;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
DB 1014 DEVLIP00GF 1023  
QY 1 DEVLIP00GF 10  
ID EGRF\_MOUSE STANDARD; PRT; 1210 AA.  
AC 001279;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Epidermal growth factor receptor precursor (EC 2.7.1.112).  
GN EGF.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c; TISSUE=Liver;  
RX MEDLINE=93026370; PubMed=1408137;  
RA Avitl A., Skorecki K., Yayon A., Givol D.;  
RT "Promoter region of the murine fibroblast growth factor receptor 2  
RT (bek/KGFR) gene.";  
RL Oncogene 7:1957-1962 (1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c, and CD-1; TISSUE=Decidua, and Liver;  
RX MEDLINE=93126380; PubMed=7678348;  
RA Paria B.C., Das S.K., Andrews G.K., Day S.K.;  
RT "Expression of the epidermal growth factor receptor gene is regulated  
RT in mouse blastocysts during delayed implantation.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:55-59 (1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c; TISSUE=Liver;  
RA Hibbs M.L.;  
RL Submitted (APR-1994) to the EMBL/Genbank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=B6/C3; TISSUE=Liver;  
RX MEDLINE=94170986; PubMed=8125255;  
RA Luetteke N.C., Phillips H.K., Qiu T.H., Copeland N.G., Earp H.S.,  
RA Jenkins N.A., Lee D.C.;  
RT "The mouse waved-2 phenotype results from a point mutation in the EGF  
RT receptor tyrosine kinase.";  
RL Gene Dev. 8:399-413 (1994).  
RN [5]  
RP SEQUENCE OF 1-714 FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=91232866; PubMed=2030916;



RA Avila A., Lax I., Ullrich A., Schlessinger J., Givol D., Morse B.;  
 RT "Comparison of EGF receptor sequences as a guide to study the ligand  
 binding site.";  
 RT Oncogene 6:673-676(1991).  
 RL [6]  
 RP SEQUENCE OF 969-1117 FROM N.A.  
 RC STRAIN-C3H;  
 RA Bisinger D.P., Serrero G.;  
 RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,  
 AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND  
 VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 tyrosine phosphate.  
 CC -1- SUBUNIT: Binds RPK1 (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- MISCELLANEOUS: Binding of EGF to the receptor leads to  
 dimerization, internalization of the EGF-receptor complex,  
 induction of the tyrosine kinase activity, stimulation of cell DNA  
 synthesis, and cell proliferation.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X78987; CAAS5587.1; -;  
 DR EMBL; U03425; AAA17899.1; -;  
 DR EMBL; X59698; CAA42219.1; -;  
 DR EMBL; L06864; AAA53029.1; -;  
 DR EMBL; Z12508; CAA78249.1; -;  
 DR PIR; A53183; A53183.  
 DR HSSP; P11362; 1FGK.  
 DR MGD; MG1:95294; Egfr.  
 DR GO; GO:0005623; Cyendocytic vesicle; IDA.  
 DR GO; GO:0030139; C:intracellular; IDA.  
 DR InterPro; IPR000494; EGFR\_L\_domain.  
 DR InterPro; IPR006211; Furin-like.  
 DR InterPro; IPR006212; Furin repeat.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF000757; Furin-like; 1.  
 DR Pfam; PF000659; Kinase; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 2.  
 DR PRINTS; PRO0109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00219; Tyrc; 1.  
 DR SMART; SM00261; Fuf; 5.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR Transmembrane; Glycoprotein; Receptor; Signal; Transferase;  
 KW Tyrosine-protein kinase; ATP-binding; Phosphorylation; Repeat.  
 FT SIGNAL 1 24  
 FT CHAIN 1 1210  
 FT DOMAIN 25 647  
 FT TRANSMEM 648 670  
 FT DOMAIN 671 1210  
 FT REPEAT 75 300  
 FT REPEAT 390 600  
 FT DOMAIN 1028 1071  
 FT DOMAIN 981 1071  
 FT NP\_BIND 714 981  
 FT BINDING 747 728  
 FT ACT\_SITE 839 839  
 FT DISULFID 190 199  
 FT DISULFID 194 207  
 FT DISULFID 215 223  
 FT DISULFID 219 231

FT DISULFID 232 240  
 FT DISULFID 236 248  
 FT DISULFID 251 260  
 FT DISULFID 264 291  
 FT DISULFID 295 307  
 FT DISULFID 311 326  
 FT DISULFID 329 333  
 FT DISULFID 329 333  
 FT DISULFID 506 515  
 FT DISULFID 510 523  
 FT DISULFID 526 535  
 FT DISULFID 539 555  
 FT DISULFID 568 571  
 FT DISULFID 562 579  
 FT DISULFID 582 591  
 FT DISULFID 595 617  
 FT DISULFID 620 628  
 FT DISULFID 624 636  
 FT MOD\_RES 680 680  
 FT MOD\_RES 1092 1092  
 FT MOD\_RES 1110 1110  
 FT MOD\_RES 1172 1172  
 FT MOD\_RES 1197 1197  
 FT CARBOHYD 128 128  
 FT CARBOHYD 175 175  
 FT CARBOHYD 196 196  
 FT CARBOHYD 352 352  
 FT CARBOHYD 413 413  
 FT CARBOHYD 444 444  
 FT CARBOHYD 528 528  
 FT CARBOHYD 568 568  
 FT CARBOHYD 603 603  
 FT CARBOHYD 623 623  
 FT CONFLICT 19 19  
 FT CONFLICT 539 539  
 FT CONFLICT 991 991  
 FT CONFLICT 1117 1117  
 FT SEQUENCE 1210 AA; 134853 MW; 690E20D46D2FD2F5 CRC64;  
 Query Match: 92.6%; Score 50; DB 1; Length 1210;  
 Best Local Similarity 80.0%; Pred. No. 0.066; Indels 0; Gaps 0;  
 Matches 8; Conservative 2; Mismatches 0;  
 Qy 1 EBYLVPQGGF 10  
 Db 1016 DEYLIPQGGF 1025  
 RESULT 7  
 ERBB\_AVIEU STANDARD; PRT; 540 AA.  
 ID ERBB\_AVIEU  
 AC P11273;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Tyrosine-protein kinase transforming protein erbB (BC 2.7.1.112).  
 GN V-ERBB.  
 OS Avian erythroblastosis virus (strain ts167).  
 OC Viruses; Retroviral viruses; Retroviridae; Avian type C retroviruses.  
 OX NCBI\_TaxID=103898;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87064458; PubMed=2878364;  
 RA Choi O.R., Trautnor C., Graf T., Beug H., Engel J.D.;  
 RT "A single amino acid substitution in v-erbB confers a thermolabile  
 phenotype to ts167 avian erythroblastosis virus-transformed erythroid  
 cells.";  
 RL Mol. Cell. Biol. 6:1751-1759(1986).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 tyrosine phosphate.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; M13179; AAA42401.1; -  
 CC PIR; A25231; TVFVEB.  
 CC HSSP; P1362; 1FGK.  
 CC InterPro; IPR000719; Prot\_kinase.  
 CC InterPro; IPR001245; Tyr\_kinase.  
 CC Pfam; PF00069; pkinase; 1.  
 CC PRINTS; PR00109; TYRKINASE.  
 CC SMART; SMO0219; Tyrc; 1.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC Transferrase; Tyrosine-protein kinase; ATP-binding; Oncogene;  
 CC Glycoprotein; Phosphorylation.  
 CC DOMAIN 132 399 PROTEIN\_KINASE.  
 CC NP\_BIND 138 146 ATP (BY SIMILARITY).  
 CC BINDING 165 165 ATP (BY SIMILARITY).  
 CC ACT\_SITE 257 257 BY SIMILARITY.  
 CC FT VARIANT 270 270 H -> D (IN THERMOLABILE V-ERBB).  
 CC SEQUENCE 540 AA; 60412 MW; 5B53297AA068B65D CRC64;

Query Match 85.2%; Score 46; DB 1; Length 540;  
 Best Local Similarity 80.0%; Pred. No. 0.19;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EBYLVPOQGF 10  
 Db 434 DEYLVPHQGF 443

RESULT 8  
 ERBB AVIER STANDARD; PRT; 604 AA.  
 ID EBYLVPOQGF 10  
 AC P00535;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Tyrosine-protein kinase transforming protein erbB (EC 2.7.1.112).  
 GN V-ERBB.  
 OS Avian erythroblastosis virus (strain BS4).  
 OC Viruses; Retrovirdae; Retroviridae; Avian type C retroviruses.  
 OX NCBI\_TaxId=79685;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H;  
 RA MEDLINE=84026539; PubMed=633222;  
 RA Yamamoto T., Nishida T., Miyajima N., Kawai S., Ooi T., Toyoshima K.;  
 RT "The erbB gene of avian erythroblastosis virus is a member of the src  
 RT gene family.";  
 RL Cell 35:71-78(1983).

RP [2]  
 RP SEQUENCE OF 1-152 FROM N.A.  
 RP MEDLINE=84223957; PubMed=6328658;  
 RA Debilitre B., Henry C., Benaisa M., Biserle G., Claverie J.-M.,  
 RA Saulle S., Martin P., Stehelin D.;  
 RT "Sequencing the erbA gene of avian erythroblastosis virus reveals a  
 RT new type of oncogene.";  
 RL Science 224:1456-1459(1984).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- DISEASE: THE V-ERBB ONCOGENE TRANSFORMS AVIAN FIBROBLASTS AND  
 CC ERYTHROBLASTS IN CULTURE AND INDUCES SARCOMA AND ERYTHROLEUKEMIAS  
 CC IN CHICKENS.  
 CC -1- MISCELLANEOUS: V-ERBB IS A TRUNCATED AND MUTATED VERSION OF THE  
 CC RECEPTOR FOR EPIDERMAL GROWTH FACTOR.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; K02006; AAA42394.1; ALT\_INIT.  
 CC PIR; A00644; TVYH.  
 CC HSSP; P1362; 1FGK.  
 CC InterPro; IPR000719; Prot\_kinase.  
 CC InterPro; IPR001245; Tyr\_kinase.  
 CC Pfam; PF00069; pkinase; 1.  
 CC PRINTS; PR00109; TYRKINASE.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC Transferrase; Tyrosine-protein kinase; ATP-binding; Oncogene;  
 CC Glycoprotein; Phosphorylation.  
 CC DOMAIN 132 399 PROTEIN\_KINASE.  
 CC NP\_BIND 138 146 ATP (BY SIMILARITY).  
 CC BINDING 165 165 ATP (BY SIMILARITY).  
 CC ACT\_SITE 257 257 BY SIMILARITY.  
 CC FT CONFLICT 29 29 R -> W (IN REF. 2).  
 CC FT CONFLICT 140 140 S -> F (IN REF. 2).  
 CC FT CONFLICT 146 146 I -> V (IN REF. 2).  
 CC SEQUENCE 604 AA; 67633 MW; 76B8CD06745D609 CRC64;

Query Match 85.2%; Score 46; DB 1; Length 604;  
 Best Local Similarity 80.0%; Pred. No. 0.21;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EBYLVPOQGF 10  
 Db 434 DEYLVPHQGF 443

RESULT 9  
 ERBB ALV STANDARD; PRT; 634 AA.  
 ID EBYLVPOQGF 10  
 AC P00534;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Tyrosine-protein kinase transforming protein erbB (EC 2.7.1.112).  
 GN V-ERBB.  
 OS Avian leukosis virus.  
 OC Viruses; Retrovirdae; Retroviridae; Alpharetrovirus.  
 OX NCBI\_TaxId=11864;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RP MEDLINE=85228222; PubMed=2988784;  
 RA Nielsen T.W., Maroney P.A., Goodwin R.G., Rottman F.M.,  
 RA Crittenden L.B., Raines M.A., Kung H.-J.;  
 RT "c-erbB activation in ALV-induced erythroblastosis: novel RNA  
 RT processing and promoter insertion result in expression of an  
 RT amino-truncated EGF receptor.";  
 RL Cell 41:719-726(1985).

CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- MISCELLANEOUS: THIS PROTEIN IS PRODUCED BY ERYTHROLEUKEMIA CELLS  
 CC IN CHICKS WHICH HAVE BEEN INJECTED WITH THE AVIAN LEUKOSIS VIRUS  
 CC AND SUBSEQUENTLY DEVELOP LEUKEMIA. THE LEUKEMIAS ARE INDUCED BY  
 CC THE INSERTION OF RETROVIRAL PROMOTER ELEMENTS IN OR NEAR THE  
 CC C-ERBB GENE, WHICH LEADS TO C-ERBB ACTIVATION.  
 CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ENV-ERBB  
 CC PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

DR EMBL; M10066; AAA48763.1; ALT\_INIT.  
DR HSSP; P11362; IFGK.  
DR InterPro; IPR000719; Prot kinase.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR Pfam; PF00069; Pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot kinase; 1.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR TRANSFERASE; Tyrosine-protein kinase; ATP-binding; Oncogene;  
KW Glycoprotein; Phosphorylation.  
FT DOMAIN 132 399 PROTEIN\_KINASE  
FT NP\_BIND 138 146 ATP (BY SIMILARITY).  
FT BINDING 165 165 ATP (BY SIMILARITY).  
FT ACT\_SITE 257 257 BY SIMILARITY.  
SQ SEQUENCE 634 AA; 70891 MW; E705E33A0BE01FCC CRC64;

Query Match 85.2%; Score 46; DB 1; Length 634;  
Best Local Similarity 80.0%; Pred. No. 0.22;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EBYLVPQGGF 10  
Db 434 DEYLVPHQGF 443

RESULT 10  
YH73 SYNY3 STANDARD; PRT; 232 AA.  
AC P73623;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical protein s111773.  
GN SLL1773.  
OS Synecocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
ON NCBI\_TaxID=1148;  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneo T., Sato S., Kocani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima T., Hirosewa M., Sugitara M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nario K., Okumura S.,  
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
CC -1- SIMILARITY: BELONGS TO THE PIRIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

DR EMBL; D90908; BAA17668.1; -.  
DR PIR; S77110; S77110.

DR InterPro; IPR007113; Cupin\_sup.  
DR InterPro; IPR003829; DUF209.  
DR Pfam; PF02678; Pirin; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 232 AA; 25700 MW; 09B9DC65352A470 CRC64;

Query Match 70.4%; Score 38; DB 1; Length 232;  
Best Local Similarity 60.0%; Pred. No. 3.2;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EBYLVPQGGF 10  
Db 45 EBYLVPQGGF 54

RESULT 11  
ERB4\_MOUSE  
ID ERB4\_MOUSE STANDARD; PRT; 95 AA.  
AC Q61527;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Receptor protein-tyrosine kinase erbB-4 (BC 2.7.1.112) (Fragments).  
GN ERB4 OR HER4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
ON NCBI\_TaxID=10090;  
RX MEDLINE=97476287; PubMed=9334263;  
RA Elenius K., Corfas G., Paul S., Choi C.J., Rio C., Plowman G.D.,  
RA Klagesbrun M.;  
RT "A novel juxtamembrane domain isoform of HER4/ErbB4. Isoform-specific  
RT tissue distribution and differential processing in response to  
RT phorbol ester.";  
RT J. Biol. Chem. 272:26761-26768(1997).  
RN [2]  
RP SEQUENCE OF 28-95 FROM N.A.  
RC SPRAIN=C57BL/6; TISSUE=Brain;  
RX MEDLINE=96069911; PubMed=7589796;  
RA Moscoso L.M., Chu G.C., Gautam M., Noakes P.G., Merlie J.P.,  
RA Sanes J.R.;  
RT "Synapse-associated expression of an acetylcholine receptor-inducing  
RT protein, ARIA/hergulin, and its putative receptors, ErbB2 and ErbB3,  
RT in developing mammalian muscle.";  
RL Dev. Biol. 172:158-169(1995).  
CC -1- FUNCTION: Specifically binds and is activated by neurotrophins, NRG-  
CC 2, NRG-3, heparin-binding EGF-like growth factor, betacellulin and  
CC NTAK. Interaction with these factors induces cell differentiation.  
CC Not activated by BGF, TGF-A, and amphiregulin (By similarity).  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphorylate.  
CC -1- SUBUNIT: Homodimer or heterodimer with each of the other ERBB  
CC receptors. Interacts with the PDZ domain of the synaptrophin SNTB2  
CC (By similarity).  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Comment=The 2 isoforms differ functionally in their response to  
CC phorbol ester: isoform JM-A is processed but not isoform JM-B.  
CC So, they respectively represent cleavable and noncleavable  
CC forms of the receptor. Both isoforms are expressed in  
CC cerebellum, cerebral cortex, spinal cord, medulla oblongata,  
CC and eye, but the kidney expresses solely isoform JM-A and the  
CC heart solely isoform JM-B;  
CC Name=JM-A;  
CC IsoId=Q61527-1; Sequence=Displayed;  
CC Name=JM-B;  
CC IsoId=Q61527-2; Sequence=VSP\_002896;  
CC -1- PTM: Ligand-binding increases phosphorylation on tyrosine  
CC residues.

```

CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: L47241; AAA93534.1;
CC DR MGD; MGJ:104771; ErbB4.
CC DR GO; GO:0045165; P:cell fate commitment; IDA.
CC DR GO; GO:0007507; P:heart development; IMP.
CC DR GO; GO:0007399; P:neurogenesis; IMP.
CC DR InterPro: IPR000719; Prot kinase.
CC DR InterPro: IPR001245; Tyr_kinase.
CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; PARTIAL.
CC DR PROSITE; PS00109; PROTEIN_KINASE_TYR; PARTIAL.
CC DR PROSITE; PS00011; PROTEIN_KINASE_DOM; PARTIAL.
CC DR Transmembrane; Glycoprotein; Multigene family; Receptor; Transferase;
CC Tyrosine-protein kinase; ATP-binding; Phosphorylation;
CC Alternative splicing.
CC FT NON_TER 1
CC FT NON_CONS 27 28 NGPTSHDCIYYPWTGSHSTLPQHA -> IGSSIEDCIQLTD
CC FT VARSPLIC 3 25 (in isoform JM-B).
CC FT /FTid=VSP_002896.
CC FT NON_TER 95 95
CC FT SEQUENCE 95 AA; 10524 MW; B43D0E99591744D8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 37; DB 1; Length 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPO 7
DB 29 EBYLVPO 35

RESULT 12
TGM2 BOVIN STANDARD; PRT; 687 AA.
AC PS1176;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Protein-glutamine gamma-glutamyltransferase (EC 2.3.2.13) (Tissue
DE transglutaminase) (TCase C) (TGC) (TG(C)) (Transglutaminase 2).
GN TGM2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCB1_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 79-95; 157-166; 242-251 AND
RP 581-587.
RC TISSUE=artery;
RX MEDLINE=92037637; PubMed=1682150;
RA Nakamishi K., Nara K., Hagiwara H., Aoyama Y., Ueno H., Hirose S.;
RT "Cloning and sequence analysis of cDNA clones for bovine aortic-
RT endothelial-cell transglutaminase.";
RL Eur. J. Biochem. 202:15-21(1991).
CC -1- FUNCTION: CATALYZES THE CROSS-LINKING OF PROTEINS AND THE
CC CONJUGATION OF POLYAMINES TO PROTEINS.
CC -1- CATALYTIC ACTIVITY: Protein glutamine + alkylamine = protein N(5)-
CC alkylglutamine + NH(3).
CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS ARE DETECTED IN THE LUNG. LOWER
CC LEVELS ARE FOUND IN THE LIVER, SPLEEN AND HEART, BUT NOT IN THE
CC BRAIN.

```

```

CC -1- INDUCTION: By retinoic acid.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: X60866; CAA43097.1;
CC DR PIR; S19680; S19680.
CC DR HSSP; P00488; IQRK.
CC DR InterPro: IPR001102; Glutransfg.
CC DR InterPro: IPR002931; Transglutaminase_like.
CC DR Pfam; PF01841; Transglut_core; 1.
CC DR Pfam; PF00927; Transglutamin_C; 2.
CC DR Pfam; PF00868; Transglutamin_N; 1.
CC DR SMART; SM00460; TGC; 1.
CC DR PROSITE; PS00547; TRANSGLUTAMINASES; 1.
CC DR Transferrase; Acyltransferase; Calcium-binding.
CC FT ACT_SITE 277 277 BY SIMILARITY.
CC FT ACT_SITE 335 335 BY SIMILARITY.
CC FT ACT_SITE 358 358 BY SIMILARITY.
CC FT METAL 398 398 CALCIUM (BY SIMILARITY).
CC FT METAL 400 400 CALCIUM (BY SIMILARITY).
CC FT METAL 447 447 CALCIUM (BY SIMILARITY).
CC FT METAL 452 452 CALCIUM (BY SIMILARITY).
CC FT SEQUENCE 687 AA; 77112 MW; 7BBA00F15E779944 CRC64;

Query Match
Best Local Similarity 60.0%; Score 37; DB 1; Length 687;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10
DB 157 QBYLVPOQGF 166

RESULT 13
TGM2 HUMAN STANDARD; PRT; 687 AA.
AC P21980; Q16436; Q9BTN7; Q9UH35;
DT 01-AUG-1991 (Rel. 19, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Protein-glutamine gamma-glutamyltransferase (EC 2.3.2.13) (Tissue
DE transglutaminase) (TCase C) (TGC) (TG(C)) (Transglutaminase 2)
DE (TCase-H).
GN TGM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RP TISSUE=Endothelial cells;
RX MEDLINE=91093168; PubMed=1670766;
RA Gentile V., Saydak M., Chiozza E.A., Akande O., Bircbachler P.J.,
RA Lee K.N., Stein J.P., Davies P.J.A.;
RT "Isolation and characterization of cDNA clones to mouse macrophage
RT and human endothelial cell tissue transglutaminases.";
RL J. Biol. Chem. 266:478-483(1991).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=93054562; PubMed=1358880;
RA Fraaij B.M., Bircbachler P.J., Patterson M.K. Jr., Lee K.N.,
RA Gonzalez R.A.;
RT "A retinoic acid-inducible mRNA from human erythroleukemia cells
RT encodes a novel tissue transglutaminase homologue.";
RL J. Biol. Chem. 267:22616-22623(1992).
RN [3]

```

RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RX MEDLINE=96201707; PubMed=8611626;  
 RA Fraj B.M., Gonzales R.A.;  
 RT "A third human tissue transglutaminase homologue as a result of  
 alternative gene transcripts";  
 RL Blochim. Biophys. Acta 1306:63-74 (1996).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Babbage A.K., Bagguley C.L.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clapp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grahman D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jokusch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Leivaesalho M.H., Leverhwa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA March V.L., Martin S.L., McComachie L.J., McLeay K., McNurray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prachinang S.R., Plumb R.W., Ramsey H.,  
 RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaundin M., Walli M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilmshurst L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20";  
 RL Nature 414:865-871 (2001).  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh L.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schreier T.E.,  
 RA Brownstein M.J., Ueda T.B., Toehlyuk S., Carninci P., Prange C.J.,  
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield Y.S.N., Krzywicki M.I., Skalek U., Smilans D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [6]  
 RP FUNCTION: CATALYZES THE CROSS-LINKING OF PROTEINS AND THE  
 CONJUGATION OF POLYAMINES TO PROTEINS.  
 CC -1- CATALYTIC ACTIVITY: Protein glutamine + alkylamine = protein N(5)-  
 alkylglutamine + NH(3).  
 CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).  
 CC -1- SUBUNIT: Monomer.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=1;  
 CC IsoId=P21980-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=P21980-2; Sequence=VSP\_006411, VSP\_006412;  
 CC Name=3; Synonyms=IGH2;

CC IsoId=P21980-3; Sequence=VSP\_006413, VSP\_006414;  
 CC -1- INDUCTION: By retinoic acid.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSGLUTAMINASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the European Bioinformatics Institute and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC  
 DR EMBL: MS5153; AAA63261.1; -;  
 DR EMBL: M98478; AAA6739.1; -;  
 DR EMBL: S81734; AAB36379.1; -;  
 DR EMBL: AL031651; CAB66115.1; -;  
 DR EMBL: AL031651; CAB66116.1; -;  
 DR EMBL: BC003551; AAH03551.1; -;  
 DR PIR: A39045; A39045.  
 DR PIR: S68092; S68092.  
 DR PDB: 1FAV; 21-JUL-00.  
 DR PDB: 1KV3; 13-MAR-02.  
 DR Gene: HGNC:11778; TGM2.  
 DR MIM: 190196; -;  
 DR GO: GO:0006464; P:protein modification; TAS.  
 DR InterPro: IPR001102; Glutaminase.  
 DR InterPro: IPR002931; Transglutaminase like.  
 DR Pfam: PF01641; Transglut core; 1.  
 DR Pfam: PF00927; Transglutamin C; 2.  
 DR Pfam: PF00868; Transglutamin N; 1.  
 DR SMART: SM00460; TSC; 1.  
 DR PROSITE: PS00547; TRANSGLUTAMINASES; 1.  
 DR Transfaser: Acyltransferase; Calcium-binding; Alternative splicing;  
 KM 3D-structure.  
 FT ACT\_SITE 277 277  
 FT ACT\_SITE 335 335  
 FT ACT\_SITE 358 358  
 FT METAL 398 398  
 FT METAL 400 400  
 FT METAL 447 447  
 FT METAL 452 452  
 FT METAL 539 548  
 FT VARSPIC 549 687  
 FT VARSPIC 287 349  
 FT VARSPIC 350 687  
 FT CONFLICT 51 51  
 FT CONFLICT 186 186  
 FT CONFLICT 224 224  
 FT CONFLICT 533 533  
 FT CONFLICT 655 655  
 FT SEQUENCE 687 AA; 77328 MM; 7DA33FF35DE7B37 CRC64;  
 Query Match 68.5%; Score 37; DB 1; Length 687;  
 Best Local Similarity 60.0%; Pred. No. 15;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EHYVPOQGF 10  
 Db 157 QEVYVPOQGF 166  
 RESULT 14  
 TGM2\_CAVCU STANDARD; PRT; 689 AA.  
 ID TGM2\_CAVCU  
 AC P08587;  
 DT 01-AUG-1988 (Rel. 08, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Protein-glutamine gamma-glutamyltransferase (EC 2.3.2.13) (Tissue  
 DE transglutaminase) (Tgase C) (TGC) (TGC(C)) (Transglutaminase 2).  
 GN TGM2.  
 OS Cavia cutleri (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystriognathi; Cavidae; Cavia.  
 OX NCBI\_TaxId=10144;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=88294033; PubMed=2900023;  
 RA Ikura K., Nasu T.-A., Yokota H., Teuchiya Y., Sasaki R., Chiba H.;  
 RT "Amino acid sequence of guinea pig liver transglutaminase from its  
 RT cDNA sequence.";  
 RL Biochemistry 27:2898-2905(1988).  
 RN [2]  
 RP SEQUENCE OF 188-631 FROM N.A.  
 RC TISSUE=Liver;  
 RA Ikura K., Nasu T.-A., Yokota H., Sasaki R., Chiba H.;  
 RT "Cloning of cDNA coding for guinea pig liver transglutaminase.";  
 RL Agric. Biol. Chem. 51:957-961(1987).  
 RN [3]  
 RP SEQUENCE OF 1-4.  
 RC TISSUE=Liver;  
 RX MEDLINE=7111415; PubMed=5543674;  
 RA Connellan J.M., Chung S.T., Whetzel N.K., Bradley L.M., Folk J.E.;  
 RT "Structural properties of guinea pig liver transglutaminase.";  
 RL J. Biol. Chem. 246:1093-1098(1971).  
 CC -1- FUNCTION: CATALYZES THE CROSS-LINKING OF PROTEINS AND THE  
 CC CONJUGATION OF POLYAMINES TO PROTEINS.  
 CC -1- CATALYTIC ACTIVITY: Protein glutamine + alkylamine = protein N(5)-  
 CC alkylglutamine + NH(3).  
 CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).  
 CC -1- SUBUNIT: Monomer.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSGLUTAMINASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M16646; AAA37056.1; -;  
 DR EMBL; D00114; BAA00068.1; -;  
 DR HSSP; P00488; ICGU.  
 DR InterPro: IPR001102; Glutransfg.  
 DR InterPro: IPR002931; Transglutaminase\_1like.  
 DR Pfam; PF01841; Transgluc\_core; 1.  
 DR Pfam; PF00927; Transglutamin\_C; 2.  
 DR Pfam; PF00868; Transglutamin\_N; 1.  
 DR SMART; SMO0460; TGC; 1.  
 DR PROSITE; PS00547; TRANSGLUTAMINASES; 1.  
 DR Transference; Acyltransferase; Calcium-binding.  
 FT INIT MET 0  
 FT MOD\_RES 1  
 FT ACT\_SITE 276  
 FT ACT\_SITE 334  
 FT ACT\_SITE 357  
 FT METAL 397  
 FT METAL 399  
 FT METAL 445  
 FT METAL 450  
 FT CONFLICT 1  
 FT CONFLICT 291  
 FT CONFLICT 335  
 FT SEQUENCE 689 AA; 77010 MW; 05E8A2794340DAF8 CRC64;  
 Query Match 68.5%; Score 37; DB 1; Length 689;

Best local similarity 60.0%; Pred. No. 15;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ERYLPPQGGF 10  
 DB 156 QERYLPPQGGF 165  
 RESULT 15  
 ERYB4\_HUMAN STANDARD; PRT; 1308 AA.  
 ID ERYB4\_HUMAN  
 AC Q15303;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112)  
 DE (p180erbB4) (Tyrosine kinase-type cell surface receptor HER4).  
 GN ERBB4 OR HER4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM JM-A).  
 RC TISSUE=Breast carcinoma;  
 RX MEDLINE=93189574; PubMed=8383326;  
 RA Plozman G.D., Culoscou J.-M., Whitney G.S., Green J.M., Carlton G.W.,  
 RA Foy L., Neubauer M.G., Shoyab M.;  
 RT "Ligand-specific activation of HER4/p180erbB4, a fourth member of the  
 RT epidermal growth factor receptor family.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:1746-1750(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS JM-A AND JM-B).  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=97476287; PubMed=9334263;  
 RA Elenius K., Cortas G., Paul S., Choi C.J., Rio C., Plozman G.D.,  
 RA Klagebun M.;  
 RT "A novel juxtamembrane domain isoform of HER4/ErbB4. Isoform-specific  
 RT tissue distribution and differential processing in response to  
 RT phorbol ester.";  
 RL J. Biol. Chem. 272:26761-26768(1997).  
 RN [3]  
 RP INTERACTION WITH SNTB2.  
 RX MEDLINE=20202681; PubMed=10725395;  
 RA Garcia R.A., Vasudevan K., Buonanno A.;  
 RT "The neurotrophin receptor ErbB-4 interacts with PDZ-containing proteins  
 RT at neuronal synapses.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:3596-3601(2000).  
 CC -1- FUNCTION: Specifically binds and is activated by neurotrophins, NRG-  
 CC 2, NRG-3, heparin-binding EGF-like growth factor, betacellulin and  
 CC NTAR.1 interaction with these factors induces cell differentiation.  
 CC Not activated by EGF, TGF- $\alpha$ , and amphiregulin.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBUNIT: Homodimer or heterodimer with each of the other ERBB  
 CC receptors (Potential). Interacts with the PDZ domain of the  
 CC synrophin SNTB2.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2.  
 CC Comment=The 2 isoforms differ functionally in their response to  
 CC phorbol ester: Isoform JM-A is processed but not isoform JM-B.  
 CC So, they respectively represent cleavable and noncleavable  
 CC forms of the receptor. Both isoforms are expressed in the  
 CC cerebellum, but only the isoform JM-B is expressed in the  
 CC heart;  
 CC Name=JM-A;  
 CC IsoId=Q15303-1; Sequence=Displayed;  
 CC Name=JM-B;  
 CC IsoId=Q15303-2; Sequence=VSP\_002895;  
 CC -1- TISSUE SPECIFICITY: Expressed at highest levels in brain, heart,  
 CC kidney, in addition to skeletal muscle, parathyroid, cerebellum,  
 CC pituitary, spleen, testis and breast. Lower levels in thymus,

CC lung, salivary gland, and pancreas.  
 CC -1- PTM: Ligand-binding increases phosphorylation on tyrosine  
 CC residue.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L07868; AAB59446.1; -  
 CC DR PIR: A47253; A47253.  
 CC DR HSP: P11362; 1FGK.  
 CC DR Genew: HGNC:3432; ERBB4.  
 CC MIM: 600543; -  
 DR GO: GO:0005887; C: integral to plasma membrane; TAS.  
 DR GO: GO:0008283; P: cell proliferation; TAS.  
 DR GO: GO:0007275; P: development; TAS.  
 DR GO: GO:0007048; P: oncogenesis; TAS.  
 DR InterPro: IPR000494; EGF\_L\_domain.  
 DR InterPro: IPR006211; Furin\_Like.  
 DR InterPro: IPR006212; Furin\_repeat.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR004019; YLP\_motif.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF00069; Kinase; 1.  
 DR Pfam: PF01030; Recep\_L\_domain; 2.  
 DR Pfam: PF02757; YLP; 2.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00261; Fu\_5.  
 DR SMART: SM00219; TyKc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KW Alternative splicing.  
 KW ACT\_SITE 843  
 FT SIGNAL 1  
 FT CHAIN 26 1308  
 FT DOMAIN 26 651  
 FT TRANSMEM 652 675  
 FT DOMAIN 676 1308  
 FT DOMAIN 186 334  
 FT DOMAIN 496 633  
 FT DOMAIN 718 985  
 FT NP\_BIND 724 732  
 FT BINDING 751 751  
 FT ACT\_SITE 843  
 FT DISULFID 189 197  
 FT DISULFID 193 205  
 FT DISULFID 213 221  
 FT DISULFID 217 229  
 FT DISULFID 230 238  
 FT DISULFID 234 246  
 FT DISULFID 249 258  
 FT DISULFID 262 289  
 FT DISULFID 293 304  
 FT DISULFID 308 323  
 FT DISULFID 326 330  
 FT DISULFID 330 330  
 FT DISULFID 503 512  
 FT DISULFID 507 520  
 FT DISULFID 523 532  
 FT DISULFID 536 552  
 FT DISULFID 555 569  
 FT DISULFID 559 577  
 FT DISULFID 580 589  
 FT DISULFID 593 614  
 FT DISULFID 617 625

FT DISULFID 621 633 BY SIMILARITY.  
 FT MOD\_RES 1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1168 1168 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1258 1258 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1284 1284 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 138 138 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 174 174 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 181 181 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 253 253 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 358 358 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 410 410 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 473 473 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 495 495 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 548 548 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 576 576 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 620 620 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT VARSPLIC 626 648 NGPTSHDCIYYPWTGSHSTLPQHA -> IGSSIDCIGLMD  
 (in isoform JM-B).  
 FT /FTId=VSP\_002895.  
 SQ SEQUENCE 1308 AA; 146807 MW; 5E4AE80985D88761 CRC64;

Query Match 68.5%; Score 37; DB 1; Length 1308;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1020 EBYLVPO 1026

Search completed: February 7, 2004, 13:26:46  
 Job time: 13 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2004, 13:26:19 ; Search time 34 Seconds  
(without alignments)  
75.898 Million cell updates/sec

Title: US-09-930-125-3

Perfect score: 54

Sequence: 1 EBYLVPQGGF 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database:

1: SP\_ARCHAEA:\*  
2: SP\_BACTERIA:\*  
3: SP\_FUNGI:\*  
4: SP\_HUMAN:\*  
5: SP\_INVERTEBRATE:\*  
6: SP\_MAMMAL:\*  
7: SP\_MHC:\*  
8: SP\_ORGANELLE:\*  
9: SP\_PHAGE:\*  
10: SP\_PLANT:\*  
11: SP\_RODENT:\*  
12: SP\_VIRUS:\*  
13: SP\_VERTEBRATE:\*  
14: SP\_UNCLASSIFIED:\*  
15: SP\_RVIRUS:\*  
16: SP\_BACTERIAP:\*  
17: SP\_ARCHAEP:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 54    | 100.0       | 367    | 11    | QBR2X1 mus musculu  |
| 2          | 54    | 100.0       | 412    | 4     | QBRVVO mus sapien   |
| 3          | 54    | 100.0       | 881    | 11    | Q8COB7 mus musculu  |
| 4          | 54    | 100.0       | 1259   | 6     | Q18735 canis famli  |
| 5          | 54    | 100.0       | 1259   | 11    | Q8K3P9 ratius famli |
| 6          | 51    | 94.4        | 1209   | 6     | Q8M1L8 sus scrofa   |
| 7          | 50    | 92.6        | 1209   | 11    | Q9GX70 mus musculu  |
| 8          | 47    | 87.0        | 1210   | 11    | Q9EP98 mus musculu  |
| 9          | 46    | 85.2        | 545    | 15    | Q85468 avian eryth  |
| 10         | 46    | 85.2        | 567    | 15    | Q86714 avian rous-  |
| 11         | 46    | 85.2        | 729    | 15    | Q86712 avian rous-  |
| 12         | 46    | 85.2        | 962    | 15    | Q64895 avian eryth  |
| 13         | 40    | 74.1        | 876    | 2     | Q32739 clostridium  |
| 14         | 40    | 74.1        | 876    | 2     | Q9KH41 clostridium  |
| 15         | 38    | 70.4        | 910    | 12    | Q9Q915 avian adeno  |
| 16         | 37    | 68.5        | 75     | 11    | Q88460 mus musculu  |

|    |    |      |      |    |        |                     |
|----|----|------|------|----|--------|---------------------|
| 17 | 37 | 68.5 | 428  | 16 | Q92BW2 | O92bw2 listeria in  |
| 18 | 37 | 68.5 | 428  | 16 | Q8Y797 | O8y797 listeria mo  |
| 19 | 37 | 68.5 | 539  | 3  | Q01143 | Q01143 magnaporthe  |
| 20 | 37 | 68.5 | 694  | 5  | Q9UB16 | Q9ub16 caenorhabdi  |
| 21 | 37 | 68.5 | 753  | 5  | Q9XZD4 | Q9xzD4 caenorhabdi  |
| 22 | 37 | 68.5 | 861  | 5  | Q8ML27 | Q8ml27 dirosophila  |
| 23 | 37 | 68.5 | 1137 | 13 | Q9W6P6 | Q9w6P6 gallus gall  |
| 24 | 36 | 66.7 | 173  | 2  | Q8GJES | Q8gJES xenorhabdus  |
| 25 | 36 | 66.7 | 332  | 4  | Q9HAA2 | Q9haa2 homo sapien  |
| 26 | 36 | 66.7 | 332  | 4  | Q9ECY5 | Q9ecy5 homo sapien  |
| 27 | 36 | 66.7 | 375  | 4  | Q8WVW8 | Q8wvW8 homo sapien  |
| 28 | 36 | 66.7 | 422  | 5  | Q9VKR8 | Q9vkr8 dirosophila  |
| 29 | 36 | 66.7 | 436  | 17 | Q97Z48 | Q97z48 sulfolobus   |
| 30 | 36 | 66.7 | 1730 | 13 | Q8UYV7 | Q8uyV7 xenopus lae  |
| 31 | 36 | 66.7 | 2259 | 11 | Q921C2 | Q921C2 mus musculu  |
| 32 | 36 | 66.7 | 2304 | 11 | Q921C3 | Q921C3 mus musculu  |
| 33 | 36 | 66.7 | 3906 | 2  | Q8G987 | Q8g987 planktothrix |
| 34 | 35 | 64.8 | 209  | 3  | Q94461 | Q94461 schizosacch  |
| 35 | 35 | 64.8 | 322  | 16 | Q8DD55 | Q8dd55 vibrio vuln  |
| 36 | 35 | 64.8 | 365  | 16 | Q987M5 | Q987M5 rhizobium l  |
| 37 | 35 | 64.8 | 481  | 13 | Q8JFP2 | Q8jfp2 brachydanio  |
| 38 | 35 | 64.8 | 516  | 3  | Q9Y2E1 | Q9y2E1 aspergillus  |
| 39 | 35 | 64.8 | 666  | 11 | Q91VG9 | Q91vg9 mus musculu  |
| 40 | 35 | 64.8 | 686  | 11 | Q9WVJ6 | Q9wvJ6 ratius norv  |
| 41 | 35 | 64.8 | 686  | 11 | Q8C217 | Q8c217 mus musculu  |
| 42 | 35 | 64.8 | 852  | 2  | Q9AM64 | Q9am64 actinobact   |
| 43 | 35 | 64.8 | 917  | 10 | Q94C76 | Q94c76 arabidopsis  |
| 44 | 35 | 64.8 | 1448 | 5  | Q8T683 | Q8t683 dictyostell  |
| 45 | 35 | 64.8 | 2269 | 4  | Q96QG9 | Q96qg9 homo sapien  |

#### ALIGNMENTS

RESULT 1  
QBR2X1 PRELIMINARY; PRT; 367 AA.

AC QBR2X1  
DT 01-JUN-2002 (TREMBLREL. 21, Created)  
DT 01-JUN-2002 (TREMBLREL. 21, Last sequence update)  
DT 01-MAR-2003 (TREMBLREL. 23, Last annotation update)

DE Hypothetical 40.2 kDa protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.  
OX NCBI\_TaxID=10090;  
[1]

RP SEQUENCE FROM N.A.

RA Straube R.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC027080; AA027080.1; .  
DR InterPro; IPR000719; Prot. Kinase.

DR InterPro; IPR004019; YLP\_Motif.

DR Pfam; PF00069; PKinase; 1.

DR Pfam; PF02757; YLP; 2.  
DR PRODOM; PD000001; Prot. Kinase; 1.

DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
KW Hypothetical protein; ATP-binding; Transferase.

SO SEQUENCE 367 AA; 40163 MW; 0BB03395F9E101B0 CRC64;

Query Match 100.0%; Score 54; DB 11; Length 367;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPQGGF 10  
DB 133 EBYLVPQGGF 142

RESULT 2

QBRVVO PRELIMINARY; PRT; 412 AA.  
ID QBRVVO  
AC QBRVVO;



```

DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypoetical protein.
GN PP3659.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang Y., Zhou X.M., Zhang P.P., Jiang H.Q., Qin W.X., Zhao X.T.,
RA Han D.F., Gu J.R.,
RT "Novel human cDNA clones with function of inhibiting cancer cell
RT growth."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF18349; AL55856.1; -
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR004019; YLP motif.
DR Pfam; PF00069; Kinase; 1.
DR Pfam; PF02757; YLP; 2.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KM Hypoetical protein; ATP-binding; Kinase; Transferase;
KM Tyrosine-protein kinase.
SQ SEQUENCE 412 AA; 44702 MW; 034397F3F27D2BC CRC64;

Query Match 100.0%; Score 54; DB 4; Length 412;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10
DB 133 EBYLVPOQGF 142

RESULT 3
Q8C0E7 PRELIMINARY; PRT; 881 AA.
ID Q8C0E7
AC Q8C0E7;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE v-erb-b2 erythroblastic leukemia viral oncogene homolog 2
DE (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Testis;
RA MEDLINE=2354683; PubMed=1246851;
RA THE FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RL EMBL; AK031542; BAC27442.1; -
FT NON TER 1
SQ SEQUENCE 881 AA; 97501 MW; 5D5042BB9F8F0836 CRC64;

Query Match 100.0%; Score 54; DB 11; Length 881;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10
DB 647 EBYLVPOQGF 656

```

```

RESULT 4
ID 018735 PRELIMINARY; PRT; 1259 AA.
AC 018735;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE ErbB-2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Placentalia; Canidae; Canis.
OC NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Yokota H.,
RT "cDNA cloning of erbB-2 from canine mammary gland."
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB008451; BAA23127.1; -
DR HSP; P11362; IFGK.
DR InterPro; IPR002048; EGFR_L_domain.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR006211; Furin-Like.
DR InterPro; IPR006212; Furin-repeat.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR004019; YLP motif.
DR Pfam; PR00757; Furin-like; 1.
DR Pfam; PF00069; Kinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR Pfam; PF02757; YLP; 2.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00261; FU; 3.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KM ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 1259 AA; 137989 MW; E37364D9C4ACD46 CRC64;

Query Match 100.0%; Score 54; DB 6; Length 1259;
Best Local Similarity 100.0%; Pred. No. 0.068;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10
DB 1020 EBYLVPOQGF 1029

RESULT 5
Q8K3F9 PRELIMINARY; PRT; 1259 AA.
ID Q8K3F9
AC Q8K3F9;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Neu proto-oncoprotein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BDIX;
RA Watson P.A., Kim K., Chen K.-S., Gould M.N.;
RT "Androgen-dependent Mammary Carcinogenesis in Rats Transgenic for the
RT Neu Proto-Oncogene."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY116182; AAWS0093.1; -
DR InterPro; IPR002048; EF-hand.

```

DR InterPro; IPR000494; EGFR\_L\_domain.  
 DR InterPro; IPR006211; Furin-like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00757; Furin-like; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 2.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TyKc; 1.  
 DR PROSITE; PS00018; EF\_HAND\_1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW ATP-binding; Transferase.  
 SQ SEQUENCE 1259 AA; 139102 MW; B724BD5CC33AE953 CRC64;

Query Match 100.0%; Score 54; DB 11; Length 1259;  
 Best Local Similarity 100.0%; Pred. No. 0.068;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EBYLVPOQGF 10  
 Db 1025 EBYLVPOQGF 1034

RESULT 6

ID Q8MIL8 PRELIMINARY; PRT; 1209 AA.  
 AC Q8MIL8;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Epidermal growth factor receptor.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.  
 OX NCBI\_Taxid=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kim J.G., Vallet J.L., Nonneman D., Christenson R.K.;  
 RT "Characterization of uterine epidermal growth factor receptor  
 expression during the estrous cycle and early pregnancy in pigs."  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY117054; AAM77472.1; -  
 DR InterPro; IPR000345; CytC\_heme\_bind.  
 DR InterPro; IPR000494; EGFR\_L\_domain.  
 DR InterPro; IPR006211; Furin-like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00757; Furin-like; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 2.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TyKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW ATP-binding; Receptor; Transferase.  
 SQ SEQUENCE 1209 AA; 133531 MW; 268E3FB11E36F90F CRC64;

Query Match 94.4%; Score 51; DB 6; Length 1209;  
 Best Local Similarity 90.0%; Pred. No. 0.26;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EBYLVPOQGF 10  
 Db 1025 EBYLVPOQGF 1034

Db 1014 DEYLVPOQGF 1023

RESULT 7

ID Q9QX70 PRELIMINARY; PRT; 1209 AA.  
 AC Q9QX70;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Epidermal growth factor receptor.  
 GN EGFR.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_Taxid=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Fisher; TISSUE=Liver;  
 RX MEDLINE=90258888; PubMed=2342466;  
 RA Petch L.A., Harris J., Raymond V.W., Blasband A.J., Lee D.C.,  
 RA Earp H.S.;  
 RT "A truncated, secreted form of the epidermal growth factor receptor is  
 encoded by an alternatively spliced transcript in normal rat tissue";  
 RL Mol. Cell. Biol. 10:2973-2982 (1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Fisher; TISSUE=Liver;  
 RA Petch L.A.;  
 RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Fisher; TISSUE=Liver;  
 RA Guttridge K., Dawson T.L., Earp H.S.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; M37394; AAF14008.1; -  
 DR HSSP; P11362; IFGK.  
 DR InterPro; IPR000494; EGFR\_L\_domain.  
 DR InterPro; IPR006212; Furin-repeat.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00757; Furin-like; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 2.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00261; FU\_3.  
 DR SMART; SM00219; TyKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW ATP-binding; Kinase; Receptor; Transferase; Tyrosine-protein kinase.  
 SQ SEQUENCE 1209 AA; 134881 MW; 96FE7F6CC1B773 CRC64;

Query Match 92.6%; Score 50; DB 11; Length 1209;  
 Best Local Similarity 80.0%; Pred. No. 0.41;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EBYLVPOQGF 10  
 Db 1015 DEYLVPOQGF 1024

RESULT 8

ID Q9EP98 PRELIMINARY; PRT; 1210 AA.  
 AC Q9EP98;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Epidermal growth factor receptor isoform 1.  
 GN EGFR.

OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/101, 129/SvJ, and 129/SvEvTAC;  
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,  
 RA Sinclair C.S., Pearshall R.S., Green P.J., Yee D., Lampland A.L.,  
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,  
 RA Maibhe N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative Egr transcripts encoding truncated receptor  
 RT isoforms.";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J;  
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,  
 RA Schehl C., Pearshall R.S., Green P.J., Yee D., Lampland A.L.,  
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,  
 RA Maibhe N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative Egr transcripts encoding truncated receptor  
 RT isoforms.";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF275364; AAG28045.1; JOINED.  
 DR EMBL; AF275365; AAG28045.1; JOINED.  
 DR EMBL; AF275367; AAG24386.1; -.  
 DR HSSP; P11362; 1FGK.  
 DR MGD; MGI:95294; Egfr.  
 DR InterPro; IPR000345; CytC heme bind.  
 DR InterPro; IPR000494; EGFR\_L\_domain.  
 DR InterPro; IPR006211; Furin-like.  
 DR InterPro; IPR006212; Furin-repeat.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00757; Furin-like; 1.  
 DR Pfam; PF00069; kinase; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 2.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00261; Fv; 3.  
 DR SMART; SM00219; TyKc; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR ATP-binding; Kinase; Receptor; Transferase; Tyrosine-protein kinase.  
 KW ATP-binding; Kinase; Receptor; Transferase; Tyrosine-protein kinase.  
 SQ SEQUENCE 1210 AA; 134840 MW; 62CD021C9E32E18 CRC64;

Query Match 87.0%; Score 47; DB 11; Length 1210;  
 Best Local Similarity 80.0%; Pred. No. 1.6;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EBYLVPQGGF 10  
 DB 1016 DEYLVPOGGF 1025

RESULT 9  
 Q85468 PRELIMINARY; PRT; 545 AA.  
 AC Q85468;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Avian Erythroblastosis virus (T834) v-erbB gene.  
 OS Avian erythroblastosis virus.  
 OC Viruses; Retroid viruses; Retroviridae; Avian type C retroviruses.  
 RN NCBI\_TaxID=11861;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=88217326; PubMed=2897102;  
 RA Scoting P., Vennartion B., Jansen M., Graf T., Beug H., Haymann M.J.;  
 RT "Common site of mutation in the erbB gene of avian erythroblastosis  
 RT virus mutants that are temperature sensitive for transformation.";  
 RL Oncogene Res. 1:265-278 (1987).  
 DR EMBL; X06943; CAA30024.1; -.  
 DR HSSP; P11362; 1FGK.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00069; kinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00219; TyKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.  
 KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.  
 SQ SEQUENCE 545 AA; 60899 MW; 14DCB8CA0F8AF4 CRC64;

Query Match 85.2%; Score 46; DB 15; Length 545;  
 Best Local Similarity 80.0%; Pred. No. 1.1;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EBYLVPQGGF 10  
 DB 439 DEYLVPOGGF 448

RESULT 10  
 Q86714 PRELIMINARY; PRT; 567 AA.  
 ID Q86714;  
 AC Q86714;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE v-erbB protein (Fragment).  
 GN v-ERBB.  
 OS Avian rous-associated virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.  
 RN NCBI\_TaxID=11950;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94203659; PubMed=8152791;  
 RA Vennartion B., Raynoscheck C., Jansson L., Doederlein G., Lhotak V.,  
 RA Johnson A., Beug H.;  
 RT "Retroviral capture of c-erbB proto-oncogene sequences: rapid  
 RT evolution of distinct viral genomes carrying mutant v-erbB genes with  
 RT different transforming capacities.";  
 RL Oncogene 9:1307-1320 (1994).  
 DR EMBL; S69372; AAC60727.1; -.  
 DR HSSP; P11362; 1FGK.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00069; kinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00219; TyKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.  
 KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.  
 FT NON TER 1  
 SQ SEQUENCE 567 AA; 63390 MW; C6D9CBA7AD725F1 CRC64;

Query Match 85.2%; Score 46; DB 15; Length 567;  
 Best Local Similarity 80.0%; Pred. No. 1.2;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EBYLVPQGGF 10  
 DB 439 DEYLVPOGGF 448

```

RESULT 11
Q86712 ID 086712 PRELIMINARY; PRT; 729 AA.
AC 086712;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Polypeptidein.
GN POLYPROTEIN.
OS Avian rous-associated virus type 1.
OC Viruses; Retrovird viruses; Retroviridae; Alpharetrovirus.
OX NCBI_TaxID=11950;
RX MEDLINE=94203659; Pubmed=8152791;
RA Vennetrom B., Raynescheck C., Jansson L., Doederlein G., Lhotak V.,
RA Johnson A., Beug H.;
RT "Retroviral capture of c-erbB proto-oncogene sequences: rapid
RT evolution of distinct viral genomes carrying mutant v-erbB genes with
RT different transforming capacities."
RL Oncogene 9:1307-1320(1994).
DR EMBL; S69372; AAC60725.1; -
DR HSSP; P03322; 1A6S.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR004028; Retro_M.
DR InterPro; IPR001245; Tyr_Pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF02813; Retro_M; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR ATP-binding; Kinase; transferase; Tyrosine-protein kinase.
KM ATP-binding; Kinase; transferase; Tyrosine-protein kinase.
SQ SEQUENCE 729 AA; 80649 MW; 84D2F6914EFD53 CRC64;

Query Match 85.2%; Score 46; DB 15; Length 729;
Best Local Similarity 80.0%; Pred. No. 1.5;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYLVPOQGF 10
Db 588 DYLVPHQGF 597

RESULT 12
Q64895 ID 064895 PRELIMINARY; PRT; 962 AA.
AC 064895;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Gag_V-erb-A-V-erb-B protein.
GN GAG_V-ERB-A-V-ERB-B.
OS Avian erythroblastosis virus.
OC Viruses; Retrovird viruses; Retroviridae; Avian type C retroviruses.
OX NCBI_TaxID=11861;
RX MEDLINE=90206603; Pubmed=1969616;
RA Brunst A., Jackson J., Bishop J.M., McCarley D.J., Schatzman R.C.;
RT "Six amino acids from the retroviral gene gag greatly enhance the
RT transforming potential of the oncogene v-erb-B."
RL Oncogene 5:15-24(1990).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR EMBL; X53209; CA336459.1; -
DR HSSP; X52211; CA336459.1; JOINED.
DR HSSP; P10828; NDL.
DR InterPro; IPR000536; Hormone_rec_1fg.
DR InterPro; IPR000719; Prot_kinase.

```

```

DR InterPro; IPR001723; Stdhmn_receptor.
DR InterPro; IPR001245; Tyr_Pkinase.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00105; Zf-C4; 1.
DR PRINTS; PR00398; STRDHOMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00219; TYKc; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR ATP-binding; DNA-binding; Kinase; Metal-binding; Nuclear protein;
KM Receptor; Transcription; Transcription regulation; Transferase;
KM Tyrosine-protein kinase; Zinc; Zinc-finger.
SQ SEQUENCE 962 AA; 108320 MW; 3C5ABD791E4B95CE CRC64;

Query Match 85.2%; Score 46; DB 15; Length 962;
Best Local Similarity 80.0%; Pred. No. 2;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYLVPOQGF 10
Db 856 DYLVPHQGF 865

RESULT 13
Q32739 ID 032739 PRELIMINARY; PRT; 876 AA.
AC 032739;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE ADP-riboseyltransferase.
GN CTRB.
OS Clostridium difficile.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1496;
RN [1]
RX STRAIN=CD196;
RC MEDLINE=97230316; Pubmed=9119480;
RA Perelle S., Gilbert M., Bourlioux P., Cortier G., Popoff M.R.;
RT "Production of a complete binary toxin (actin-specific ADP-
RT ribosyltransferase) by Clostridium difficile CD196."
RL Infect. Immun. 65:1402-1407(1997).
DR EMBL; L76081; AAB67305.1; -
DR HSSP; P13423; IACC.
DR InterPro; IPR003896; Anthrax_toxinB.
DR Pfam; PF03495; Binary_tox; 1.
DR PRINTS; PR01391; BINARYTOXINB.
KM TRANSFERASE.
SQ SEQUENCE 876 AA; 98797 MW; 25E06E2D45CE2B3B CRC64;

Query Match 74.1%; Score 40; DB 2; Length 876;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 EYLVPOQGF 10
Db 696 DYLVPHQGF 704

RESULT 14
Q9KH41 ID 09KH41 PRELIMINARY; PRT; 876 AA.

```

AC 09K41,  
 DT 01-OCT-2000 (TEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)  
 DE CdtB.  
 GN CDTB.  
 OS Clostridium difficile.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1496;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CCUG 20309;  
 RA Chang S.Y., Song K.P.;  
 RT "ADP-ribosylating Binary Toxin Genes of Clostridium difficile strain  
 RT CCUG 20309."  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF271719; AAF81761.1; -.  
 DR HSSP; P13423; IACC.  
 DR InterPro; IPR003896; Anthrax\_toxinB.  
 DR Pfam; PF03495; Binary\_toxB; 1.  
 DR PRINTS; PR01391; BINARYTOXINB.  
 SO SEQUENCE 876 AA; 98793 MW; 366D62F352E745A5 CRC64;

Query Match 74.1%; Score 40; DB 2; Length 876;  
 Best Local Similarity 66.7%; Pred. No. 30;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 EBYLVPOQGF 10  
 DB 696 DYLVPEQGY 704

## RESULT 15

09Q915 PRELIMINARY; PRT; 910 AA.  
 AC 09Q915;  
 DT 01-MAY-2000 (TEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)  
 DE 100 kDa protein homolog (Fragment).  
 OS Avian adenovirus type 8 (Strain ATCC A-2A) (Fowl adenovirus 8).  
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadenovirus.  
 OX NCBI\_TaxID=66295;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CFA40;  
 RA Johnson M.A., Pooley C.;  
 RT "Fowl adenovirus serotype 8 hypervirulent strain CFA40."  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF155911; AAF17336.1; -.  
 DR InterPro; IPR003381; Adeno\_100.  
 DR Pfam; PF02438; adeno\_100; 1.  
 FT NON\_TER  
 SQ SEQUENCE 910 AA; 102772 MW; EB0BA4227ED2CEC CRC64;

Query Match 70.4%; Score 38; DB 12; Length 910;  
 Best Local Similarity 66.7%; Pred. No. 78;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQG 9  
 DB 855 EBYLVPEQG 863

Search completed: February 7, 2004, 13:32:06  
 Job time : 37 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2004, 13:26:49 ; Search time 22 Seconds  
(without alignments)  
19.232 Million cell updates/sec

Title: US-09-930-125-3

Sequence: 1 EBYLVPOGGF 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Length | ID                | Description       |
|------------|-------|--------------|-------------------|-------------------|
| 1          | 54    | 100.0        | US-08-414-417B-69 | Sequence 69, Appl |
| 2          | 54    | 100.0        | US-08-486-348A-69 | Sequence 69, Appl |
| 3          | 54    | 100.0        | US-08-468-545B-69 | Sequence 69, Appl |
| 4          | 54    | 100.0        | US-08-466-680B-69 | Sequence 69, Appl |
| 5          | 54    | 100.0        | US-08-467-083-69  | Sequence 68, Appl |
| 6          | 54    | 100.0        | US-08-414-417A-68 | Sequence 68, Appl |
| 7          | 54    | 100.0        | US-08-484-438-8   | Sequence 8, Appl  |
| 8          | 54    | 100.0        | US-08-486-348A-68 | Sequence 68, Appl |
| 9          | 54    | 100.0        | US-08-625-101-2   | Sequence 2, Appl  |
| 10         | 54    | 100.0        | US-08-468-545B-68 | Sequence 68, Appl |
| 11         | 54    | 100.0        | US-08-356-786-2   | Sequence 2, Appl  |
| 12         | 54    | 100.0        | US-08-466-680B-68 | Sequence 68, Appl |
| 13         | 54    | 100.0        | US-09-527-487-2   | Sequence 2, Appl  |
| 14         | 50    | 92.6         | US-08-128-971B-12 | Sequence 12, Appl |
| 15         | 50    | 92.6         | US-08-484-438-7   | Sequence 7, Appl  |
| 16         | 50    | 92.6         | US-08-475-035-4   | Sequence 4, Appl  |
| 17         | 42    | 77.8         | US-08-408-604A-69 | Sequence 69, Appl |
| 18         | 37    | 68.5         | US-09-356-818A-2  | Sequence 2, Appl  |
| 19         | 37    | 68.5         | US-08-484-438-6   | Sequence 6, Appl  |
| 20         | 37    | 68.5         | US-08-247-902A-2  | Sequence 2, Appl  |
| 21         | 37    | 68.5         | PCT-US93-10541-2  | Sequence 2, Appl  |
| 22         | 37    | 68.5         | PCT-US91-09784-2  | Sequence 2, Appl  |
| 23         | 37    | 68.5         | US-08-456-647B-4  | Sequence 4, Appl  |
| 24         | 37    | 68.5         | US-08-237-401A-4  | Sequence 4, Appl  |
| 25         | 37    | 68.5         | US-08-484-438-4   | Sequence 4, Appl  |
| 26         | 37    | 68.5         | US-08-484-438-2   | Sequence 2, Appl  |
| 27         | 36    | 66.7         | PCT-US93-01669-56 | Sequence 56, Appl |

ALIGNMENTS

|    |    |      |     |   |                      |                   |
|----|----|------|-----|---|----------------------|-------------------|
| 28 | 36 | 66.7 | 209 | 4 | US-09-252-991A-30648 | Sequence 30648, A |
| 29 | 35 | 64.8 | 516 | 4 | US-09-215-694-16     | Sequence 16, Appl |
| 30 | 35 | 64.8 | 685 | 5 | PCT-US91-09784-4     | Sequence 4, Appl  |
| 31 | 34 | 63.0 | 182 | 4 | US-09-328-352-8193   | Sequence 8193, Ap |
| 32 | 34 | 63.0 | 590 | 4 | US-09-252-991A-19046 | Sequence 19046, A |
| 33 | 33 | 61.1 | 9   | 1 | US-08-178-570-63     | Sequence 63, Appl |
| 34 | 33 | 61.1 | 9   | 3 | US-08-369-643-63     | Sequence 63, Appl |
| 35 | 33 | 61.1 | 9   | 5 | PCT-US95-00147-63    | Sequence 63, Appl |
| 36 | 33 | 61.1 | 12  | 1 | US-08-406-192-1      | Sequence 1, Appl  |
| 37 | 33 | 61.1 | 12  | 2 | US-08-545-151-1      | Sequence 1, Appl  |
| 38 | 33 | 61.1 | 12  | 6 | US-09-933-12         | Sequence 12, Appl |
| 39 | 33 | 61.1 | 27  | 6 | US-09-252-991A-27127 | Sequence 27127, A |
| 40 | 33 | 61.1 | 372 | 2 | US-08-821-355A-5     | Sequence 5, Appl  |
| 41 | 33 | 61.1 | 442 | 2 | US-09-003-687A-5     | Sequence 5, Appl  |
| 42 | 33 | 61.1 | 442 | 3 | US-09-136-605-5      | Sequence 5, Appl  |
| 43 | 33 | 61.1 | 491 | 4 | US-09-252-991A-29041 | Sequence 29041, A |
| 44 | 33 | 61.1 | 581 | 2 | US-08-724-394A-2     | Sequence 2, Appl  |
| 45 | 33 | 61.1 | 581 | 2 | US-08-724-394A-2     | Sequence 2, Appl  |

RESULT 1  
US-08-414-417B-69  
Sequence 69, Application US/08414417B  
Patent No. 5801005  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disis, Mary L.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/414,417B  
FILING DATE: 31-MAR-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 69:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 580 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-414-417B-69  
Query Match 100.0%; Score 54; DB 1; Length 580;  
Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EBYLVPOGGF 10  
DB 346 EBYLVPOGGF 355

```
RESULT 2
US-08-486-348A-69
; Sequence 69, Application US/08468348A
; Patent No. 5846538
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,348A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 580 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-486-348A-69

Query Match      100.0%; Score 54; DB 2; Length 580;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EBYLVPOQGF 10
        |||||
        346 EBYLVPOQGF 355

RESULT 3
US-08-468-545B-69
; Sequence 69, Application US/08468545B
; Patent No. 5876712
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,545B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 580 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-468-545B-69

Query Match      100.0%; Score 54; DB 2; Length 580;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,545B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 580 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-468-545B-69

Query Match      100.0%; Score 54; DB 2; Length 580;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EBYLVPOQGF 10
        |||||
        346 EBYLVPOQGF 355

RESULT 4
US-08-466-680B-69
; Sequence 69, Application US/08466680B
; Patent No. 6075122
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,680B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 580 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-466-680B-69

Query Match      100.0%; Score 54; DB 3; Length 580;
```



Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EBYLVPOOGF 10  
Db 346 EBYLVPOOGF 355

## RESULT 5

US-08-467-083-68  
Sequence 68, Application US/08467083  
Patent No. 5726023  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disla, Mary L.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN  
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
TITLE OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,083  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/414,417  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
TELEX: 3723836 SEDDANBERRY  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-467-083-68

Query Match 100.0%; Score 54; DB 1; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0.066;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EBYLVPOOGF 10  
Db 1021 EBYLVPOOGF 1030

RESULT 6  
US-08-414-417B-68  
Sequence 68, Application US/0841417B  
Patent No. 5801005  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disla, Mary L.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN  
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
TITLE OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED  
NUMBER OF SEQUENCES: 69

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/414,417B  
FILING DATE: 31-MAR-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-414-417B-68

Query Match 100.0%; Score 54; DB 1; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0.066;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EBYLVPOOGF 10  
Db 1021 EBYLVPOOGF 1030

## RESULT 7

US-08-484-438-8  
Sequence 8, Application US/08484438  
Patent No. 5811098  
Patent No. 5811098 5780031  
GENERAL INFORMATION:  
APPLICANT: Plowman, Gregory D.  
APPLICANT: Culouscou, Jean-Michel  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Siegall, Clay B.  
APPLICANT: Hellstr m, Ingegard  
APPLICANT: Hellstr m, Karl B.  
TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,438  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/323,442  
FILING DATE: 14-OCT-1994  
APPLICATION NUMBER: US 08/150,704

FILING DATE: 10-NOV-1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/981,165  
FILING DATE: 24-NOV-1992  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Mierock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-230  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-486-438-8

Query Match 100.0%; Score 54; DB 2; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0.066;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEVLVPOOGF 10  
Db 1021 EEVLVPOOGF 1030

RESULT 8  
US-08-486-348A-68  
Sequence 68, Application US/08486348A  
Patent No. 5846538  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disla, Mary L.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
NUMBER OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/486.348A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-486-348A-68

Query Match 100.0%; Score 54; DB 2; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0.066;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEVLVPOOGF 10  
Db 1021 EEVLVPOOGF 1030

RESULT 9  
US-08-625-101-2  
Sequence 2, Application US/08625101  
Patent No. 5869445  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disla, Mary L.  
TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE  
TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION  
TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/625.101  
FILING DATE: 01-APR-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C7  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-625-101-2

Query Match 100.0%; Score 54; DB 2; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0.066;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEVLVPOOGF 10  
Db 1021 EEVLVPOOGF 1030

RESULT 10  
US-08-468-545B-68  
Sequence 68, Application US/08468545B  
Patent No. 5876712  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disla, Mary L.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,545B  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-466-545B-66

Query Match 100.0%; Score 54; DB 2; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0.066;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10  
DB 1021 EBYLVPOQGF 1030

RESULT 11  
US-08-356-786-2  
Sequence 2, Application US/08356786  
Patent No. 5877305  
GENERAL INFORMATION:  
APPLICANT: Hueston, James S.  
APPLICANT: Oppermann, Hermann  
APPLICANT: Houston, L. L.  
APPLICANT: Ring, David B.  
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
TITLE OF INVENTION: Marker  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Edmund R. Pitcher, Teesta, Hurwitz, & Thibault  
STREET: Exchange Place, 53 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/356,786  
FILING DATE: 06-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/831,967  
FILING DATE: 06-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-053

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-356-786-2

Query Match 100.0%; Score 54; DB 2; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0.066;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10  
DB 1021 EBYLVPOQGF 1030

RESULT 12  
US-08-466-680B-68  
Sequence 68, Application US/08466680B  
Patent No. 6075122  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disis, Mary L.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,680B  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-466-680B-68

Query Match 100.0%; Score 54; DB 3; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0.066;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10  
DB 1021 EBYLVPOQGF 1030

RESULT 13  
US-09-527-487-2  
Sequence 2, Application US/09527487

Patent No. 6528060  
GENERAL INFORMATION:  
APPLICANT: Nicolette, Charles  
TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES  
FILE REFERENCE: 126881309200  
CURRENT APPLICATION NUMBER: US/09/527,487  
CURRENT FILING DATE: 2000-03-16  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 1255  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-527-487-2

Query Match 100.0%; Score 54; DB 4; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 0.066;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10  
DB 1021 EBYLVPOQGF 1030

RESULT 14  
US-08-128-971B-12  
Sequence 12, Application US/08128971B  
Patent No. 5525503  
GENERAL INFORMATION:  
APPLICANT: Christopher E. Rudd  
TITLE OF INVENTION: SIGNAL TRANSDUCTION VIA CD28  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
ZIP: 02110-2804  
COUNTRY: U.S.A.  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/128,971B  
FILING DATE: September 28, 1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Janis K. Frazer  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 00530/073001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-128-971B-12

Query Match 92.6%; Score 50; DB 1; Length 13;  
Best Local Similarity 80.0%; Pred. No. 0.0027;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10

DB 3 DEYLVPOQGF 12

RESULT 15  
US-08-484-438-7  
Sequence 7, Application US/08484438  
Patent No. 5811098  
GENERAL INFORMATION:  
APPLICANT: Plowman, Gregory D.  
APPLICANT: Culouscou, Jean-Michel  
APPLICANT: Shoyab, Mohamed  
APPLICANT: Siegfalt, Clay B.  
APPLICANT: Hellset m, Ingegerd  
TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Fennle & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,438  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/323,442  
FILING DATE: 14-OCT-1994  
APPLICATION NUMBER: US 08/150,704  
FILING DATE: 10-NOV-1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/981,165  
FILING DATE: 24-NOV-1992  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-230  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1210 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-484-438-7

Query Match 92.6%; Score 50; DB 2; Length 1210;  
Best Local Similarity 80.0%; Pred. No. 0.36;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10  
DB 1014 DEYLVPOQGF 1023

Search completed: February 7, 2004, 13:32:39  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: February 9, 2004, 16:12:17 ; Search time 45 Seconds  
(without alignments)  
7196.800 Million cell updates/sec

Title: US-09-930-125-2

Sequence: 1 METALCRMGLLALLPGAA.....TFKGTPTAENPEYGLDVPV 1255

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 6304   | 92.5        | 1259   | 6     | 018735      |
| 2          | 6007.5 | 88.2        | 1259   | 11    | 08K3P9      |
| 3          | 4216   | 61.9        | 881    | 11    | 08COE7      |
| 4          | 3172   | 46.5        | 1209   | 11    | 09QX70      |
| 5          | 3152.5 | 46.3        | 1209   | 6     | 08MIL8      |
| 6          | 3143   | 46.1        | 1210   | 11    | 09BP98      |
| 7          | 2758   | 40.5        | 1165   | 13    | 09YH40      |
| 8          | 2735.5 | 40.1        | 1137   | 13    | 09W6F6      |
| 9          | 2326   | 34.1        | 1328   | 13    | P79754      |
| 10         | 2209.5 | 34.4        | 1305   | 13    | 08A81       |
| 11         | 2043.5 | 30.0        | 1433   | 5     | 09B1H9      |
| 12         | 2013.5 | 29.5        | 1377   | 5     | 08MLN0      |
| 13         | 1874.5 | 27.5        | 419    | 4     | 09UK79      |
| 14         | 1747   | 25.6        | 367    | 11    | 08R2X1      |
| 15         | 1719   | 25.2        | 729    | 15    | 086712      |
| 16         | 1717   | 25.2        | 567    | 15    | 086714      |

|    |        |      |      |    |        |                     |
|----|--------|------|------|----|--------|---------------------|
| 17 | 1705.5 | 25.0 | 412  | 4  | 08WYV0 | 08WYV0 homo sapien  |
| 18 | 1651.5 | 24.2 | 962  | 15 | 064895 | 064895 avian eryth  |
| 19 | 1642   | 24.1 | 545  | 15 | 085468 | 085468 avian eryth  |
| 20 | 1533.5 | 22.5 | 655  | 11 | 09WVF5 | 09WVF5 mus musculus |
| 21 | 1517.5 | 22.3 | 643  | 11 | 09ERV6 | 09ERV6 mus musculus |
| 22 | 1286   | 18.9 | 1193 | 5  | 09YIX8 | 09YIX8 ephydactia f |
| 23 | 1197.5 | 17.6 | 1368 | 5  | 023821 | 023821 caenorhabdi  |
| 24 | 1177   | 16.3 | 1717 | 5  | 026566 | 026566 echinosoma   |
| 25 | 1155   | 16.9 | 527  | 13 | 090836 | 090836 gallus gall  |
| 26 | 1028.5 | 15.1 | 478  | 11 | 09PSH0 | 09PSH0 ractus norv  |
| 27 | 971.5  | 14.3 | 599  | 13 | 09PSH2 | 09PSH2 gallus gall  |
| 28 | 901    | 13.2 | 165  | 4  | 014256 | 014256 homo sapien  |
| 29 | 887    | 13.0 | 176  | 11 | 0923V5 | 0923V5 ractus norv  |
| 30 | 814.5  | 12.0 | 346  | 13 | P11776 | P11776 xiphophorus  |
| 31 | 778    | 11.4 | 435  | 5  | 08SZM1 | 08SZM1 dirosophila  |
| 32 | 757    | 11.1 | 311  | 13 | 099162 | 099162 xiphophorus  |
| 33 | 754.5  | 11.1 | 1362 | 13 | 09PVZ4 | 09PVZ4 xenopus lae  |
| 34 | 737    | 10.8 | 1671 | 5  | 09N1V5 | 09N1V5 biophalari   |
| 35 | 734    | 10.8 | 331  | 4  | 09BDU7 | 09BDU7 homo sapien  |
| 36 | 723    | 10.6 | 149  | 6  | 09BG66 | 09BG66 oryctolagus  |
| 37 | 704.5  | 10.3 | 1368 | 13 | 08UW85 | 08UW85 paralicthy   |
| 38 | 700    | 10.3 | 1418 | 13 | 093457 | 093457 scophthalmu  |
| 39 | 683.5  | 10.0 | 1369 | 13 | 08UW86 | 08UW86 paralicthy   |
| 40 | 680    | 10.0 | 1358 | 13 | 073798 | 073798 xenopus lae  |
| 41 | 667.5  | 9.8  | 1472 | 5  | 09U5A8 | 09U5A8 bombyx mori  |
| 42 | 659    | 9.7  | 1412 | 13 | 08UW84 | 08UW84 paralicthy   |
| 43 | 646.5  | 9.5  | 1418 | 13 | 08UW83 | 08UW83 paralicthy   |
| 44 | 646    | 9.5  | 1245 | 13 | 09YGH8 | 09YGH8 scophthalmu  |
| 45 | 642    | 9.4  | 1371 | 11 | 09QVW4 | 09QVW4 ractus sp.   |

## ALIGNMENTS

### RESULT 1

018735 ID 018735 PRELIMINARY; PRT; 1259 AA.

AC 018735; 01-JAN-1998 (TRENBLREL. 05, Created)

DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)

DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)

DE ExB-2.

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI\_TaxID=9615;

RN [1]

RP SEQUENCE FROM N.A.

RA Yokota H.;

RT "cDNA cloning of exB-2 from canine mammary gland.";

RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB008451; BAA23127.1; -

DR HSSP; P11362; IFGK.

DR InterPro; IPR002048; EF-hand.

DR InterPro; IPR000494; EGFR\_L domain.

DR InterPro; IPR006211; Furin-like.

DR InterPro; IPR006212; Furin repeat.

DR InterPro; IPR000719; Prot\_Kinase.

DR InterPro; IPR001245; Tyr\_Kinase.

DR InterPro; IPR004019; YLP\_motif.

DR Pfam; PF00757; Furin-like; 1.

DR Pfam; PF00069; Pkinase; 1.

DR Pfam; PF01030; Recep\_L\_domain; 2.

DR Pfam; PF02757; YLP\_2.

DR PRINTS; PR00109; TYRKINASE.

DR PRODOM; PD000001; Prot\_Kinase; 1.

DR SMART; SMO0261; FU; 3.

DR SMART; SMO0219; TYKC; 1.

DR PROSITE; PS00018; EF\_HAND; 1.

DR PROSITE; PS00107; PROTEIN\_KINASE\_DOM; 1.

DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.

DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.

DR ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.

SQ SEQUENCE 1259 AA; 137988 MW; E37364D49C4ACD46 CRC64;  
 Query Match 92.5%; Score 6304; DB 6; Length 1259;  
 Best Local Similarity 92.2%; Pred. No. 0;  
 Matches 1162; Conservative 38; Mismatches 54; Indels 6; Gaps 2;

```

QY 1 MELALCRWGLLALLPPGAASSTOVCCTGDMKRLPASPEETHLMDLRHLYGCGVQVGNL 60
DB 1 MELAMCRWGLLALLPPGAAGTCVCTGDMKRLPASPEETHLMDLRHLYGCGVQVGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVGVYLIANNOYRQVPLQRLIRVGTQGFEDNYALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIOEVGVYLIANNOYRQVPLQRLIRVGTQGFEDNYALAVLNG 120
QY 121 DPLNNTTPTVTVGASPGELQRLSLTEILKSGVLIQNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTTPTVTVGASPGELQRLSLTEILKSGVLIQNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLIDTNRSAKCHRCSPKCKSRCKWSESDCOSITRTVCAGGACRCAGPLPTDCHEQC 240
DB 181 LTLIDTNRSAKCHRCSPKCKSRCKWSESDCOSITRTVCAGGACRCAGPLPTDCHEQC 240
QY 241 AAGCTGPGHSDCLCLAHNHSIGICELHCPALVTYNTDFESMPNDEGRYTGASCVTACP 300
DB 241 AAGCTGPGHSDCLCLAHNHSIGICELHCPALVTYNTDFESMPNDEGRYTGASCVTACP 300
QY 301 YNYLSTDVGSCITVCPILNOEVTABDGTQCEKSKPCARVCYGLGMEHLREVAVTSAN 360
DB 301 YNYLSTDVGSCITVCPILNOEVTABDGTQCEKSKPCARVCYGLGMEHLREVAVTSAN 360
QY 361 IQEPAGCKKIFGSLAFPLESPFDGPASNTAPLOEPOLQVETLEITGYLTSAMPDLSL 420
DB 361 IQEPAGCKKIFGSLAFPLESPFDGPASNTAPLOEPOLQVETLEITGYLTSAMPDLSL 420
QY 421 DLSFONLOVTRGRILHNGAYSLTQAGISMLGSLRSLREISGLALHNTHLCFYVTV 480
DB 421 DLSFONLOVTRGRILHNGAYSLTQAGISMLGSLRSLREISGLALHNTHLCFYVTV 480
QY 481 PMDQLFRNPQALHTANRPEDECEVGEIACHQLCARHCGMPGTQCVNCSQPLRGEC 540
DB 481 PMDQLFRNPQALHTANRPEDECEVGEIACHQLCARHCGMPGTQCVNCSQPLRGEC 540
QY 541 VEECRVLTQGLFREYVNAHCLPCHEPCOPONGSVTCFEPEDQCYACAHYDPPFCVARC 600
DB 541 VEECRVLTQGLFREYVNAHCLPCHEPCOPONGSVTCFEPEDQCYACAHYDPPFCVARC 600
QY 601 PSQVAPDLSVMPIMKPEDEBACOPRCINCHSCVDLDDCKCPAEORASPLTSTISAVG 660
DB 601 PSQVAPDLSVMPIMKPEDEBACOPRCINCHSCVDLDDCKCPAEORASPLTSTISAVG 660
QY 661 ILLVVLGVVFGILIKRQOKIKRTYMRLLQETELVEPLTPSGAMPQAQMRILKETEL 720
DB 661 ILLVVLGVVFGILIKRQOKIKRTYMRLLQETELVEPLTPSGAMPQAQMRILKETEL 720
QY 721 RKVKVLSGSAFGTYKGIWIPDGENVKI PVAIKVLRNTSPKANKELIDEAYVMAVGSP 780
DB 721 RKVKVLSGSAFGTYKGIWIPDGENVKI PVAIKVLRNTSPKANKELIDEAYVMAVGSP 780
QY 781 YVSRLLGICLTSTQVLTQMLPYGCLLDHVRNRRGLSDOLLNMCQIAGMSTLEVR 840
DB 781 YVSRLLGICLTSTQVLTQMLPYGCLLDHVRNRRGLSDOLLNMCQIAGMSTLEVR 840
QY 841 LVHRDLAARNVLYVSPNHTITDFGLARLLDIDETEHADGGKIVIKMMALLESILRRFT 900
DB 841 LVHRDLAARNVLYVSPNHTITDFGLARLLDIDETEHADGGKIVIKMMALLESILRRFT 900
QY 901 HOSDWSVIGTVVWELMTFGAKPYDGI PARERIPDLLEKGERLPQPICTIDVYMIWVCKM 960
DB 901 HOSDWSVIGTVVWELMTFGAKPYDGI PARERIPDLLEKGERLPQPICTIDVYMIWVCKM 960
QY 961 IDSECRPFRRELVAEFSSMAADPQRFVYI QNEDLGPAPBLSTPFRSLLEDDDKDLVDA 1020
DB 961 IDSECRPFRRELVAEFSSMAADPQRFVYI QNEDLGPAPBLSTPFRSLLEDDDKDLVDA 1020
QY 1021 BEYLVPQGFECPPDPAAGCMVHRRSSSTRSGGDLTGLPSEBEARSPSLAPSEG 1080
DB 1021 BEYLVPQGFECPPDPAAGCMVHRRSSSTRSGGDLTGLPSEBEARSPSLAPSEG 1080
QY 1081 AGSDVFPDGDGMGAAGLQSLPQSDPPLQRYSEDPTVPLPEPFDGKVALPTCSPOPEYV 1140
DB 1081 AGSDVFPDGDGMGAAGLQSLPQSDPPLQRYSEDPTVPLPEPFDGKVALPTCSPOPEYV 1140
QY 1141 NOPDVRQPPSPRPGPLPAARPAQATLER-----PKTLSPQKNGVVKQVFAFGAVENPE 1195
DB 1141 NOPDVRQPPSPRPGPLPAARPAQATLER-----PKTLSPQKNGVVKQVFAFGAVENPE 1195
QY 1196 YLPFGGAAPQHPHPAPPAFDNLYYWDODPPERGAPPSPTFGPTAENBEYGLADVY 1255
DB 1196 YLPFGGAAPQHPHPAPPAFDNLYYWDODPPERGAPPSPTFGPTAENBEYGLADVY 1255
QY 1200 YLAPRGAAPQHPHPAPPAFDNLYYWDODPPERGAPPSPTFGPTAENBEYGLADVY 1259
DB 1200 YLAPRGAAPQHPHPAPPAFDNLYYWDODPPERGAPPSPTFGPTAENBEYGLADVY 1259
  
```

RESULT 2  
 Q8K3F9 PRELIMINARY; PRT; 1259 AA.  
 ID Q8K3F9  
 AC Q8K3F9;  
 DT 01-OCT-2002 (Tremblrel. 22, Created)  
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE New protooncoprotein.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxId=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BDIX;  
 RA Watson P.A., Kim K., Chen K.-S., Gould M.N.;  
 RT "Androgen-Dependent Mammary Carcinogenesis in Rats Transgenic for the  
 RT New Proto-Oncogene.";  
 RL Submitted (May-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AY116182; AA050093.1; -;  
 DR InterPro; IPR002048; EF-hand.  
 DR InterPro; IPR000494; EGFR\_L domain.  
 DR InterPro; IPR006211; Furin-like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_Thr\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00757; Furin-like; 1.  
 DR Pfam; PF00069; Kinase; 1.  
 DR Pfam; PF01030; Reced\_L domain; 2.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TYKc; 1.  
 DR PROSITE; PS00018; EF\_HAND; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS01019; PROTEIN\_KINASE\_TYR; 1.  
 DR ATP-binding; Transferrase.  
 SQ SEQUENCE 1259 AA; 139102 MW; B724BD5CC3A8953 CRC64;

Query Match 88.2%; Score 6007.5; DB 11; Length 1259;  
 Best Local Similarity 88.1%; Pred. No. 0;  
 Matches 1106; Conservative 49; Mismatches 100; Indels 1; Gaps 1;

```

QY 1 MELALCRWGLLALLPPGAASSTOVCCTGDMKRLPASPEETHLMDLRHLYGCGVQVGNL 60
DB 4 MELAMCRWGLLALLPPGAAGTCVCTGDMKRLPASPEETHLMDLRHLYGCGVQVGNL 63
QY 61 ELTYLPTNASLSFLQDIOEVGVYLIANNOYRQVPLQRLIRVGTQGFEDNYALAVLNG 120
DB 64 ELTYLPTNASLSFLQDIOEVGVYLIANNOYRQVPLQRLIRVGTQGFEDNYALAVLNG 123
QY 121 DPLNNTTPTVTVGASPGELQRLSLTEILKSGVLIQNPOLCYODTILMKDIFHKNNOLA 179
DB 124 DPLNNTTPTVTVGASPGELQRLSLTEILKSGVLIQNPOLCYODTILMKDIFHKNNOLA 183
  
```

```

QY 180 ALTLIDTNRSRACHPCSPKCSRWKSGESSEBDQSLTRTWACAGCARCKGRLPTDCHEQ 239
DB 184 APVIDIDTNRSRACHPCSPKCDKNHCKGESSEBDQSLTRTWACAGCARCKGRLPTDCHEQ 243
QY 240 CAGCGTGPBKSDCLACHFNHSGICEAHCPALVTYNTDTESMNPNBGRYTFGASCTAC 299
DB 244 CAGCGTGPBKSDCLACHFNHSGICEAHCPALVTYNTDTESMNPNBGRYTFGASCTAC 303
QY 300 PNYVLSTGVSCITVCPLEHNOEVTABDGTORCEKSKPCARVCYGLMEHLREAVATSA 359
DB 304 PNYVLSTGVSCITVCPLEHNOEVTABDGTORCEKSKPCARVCYGLMEHLREAVATSA 363
QY 360 NIGEPACCKKIFGSLAPLPSFPDQPSANTAPLOPEOLQVETLEETGYLYISAMPDSL 419
DB 364 NIGEPACCKKIFGSLAPLPSFPDQPSANTAPLOPEOLQVETLEETGYLYISAMPDSL 423
QY 420 PDLSPVQNLQVITGRILHNGAYSLTLOGLSIWLGLSLBELSGALIIHNTHLCFVHT 479
DB 424 RDLSPVQNLQVITGRILHNGAYSLTLOGLSIWLGLSLBELSGALIIHNTHLCFVHT 483
QY 480 VPMDOFLRNPHOALLHTANRPEDECVBGLACHOLCARGHGMGPQCVCNCSQFLRGOE 539
DB 484 VPMDOFLRNPHOALLHTANRPEDECVBGLACHOLCARGHGMGPQCVCNCSQFLRGOE 543
QY 540 CVEBECRYLQGLPREYVNAHCLPCHPECPONGSVTCFGEADQVCACAHYKDPFCVAR 599
DB 544 CVEBECRYLQGLPREYVNAHCLPCHPECPONGSVTCFGEADQVCACAHYKDPFCVAR 603
QY 600 CPBGVCKPDLSTYMPWKFPDEBGAOCPCPINCTHSCVDLDDKCPABORASPLTISIAV 659
DB 604 CPBGVCKPDLSTYMPWKFPDEBGAOCPCPINCTHSCVDLDDKCPABORASPLTISIAV 663
QY 660 GILLVVLVGVFGLILRROOKIRKYMTRLLOTELVEPLTPSGAMPNOAMILKETE 719
DB 664 GILLVVLVGVFGLILRROOKIRKYMTRLLOTELVEPLTPSGAMPNOAMILKETE 723
QY 720 LRKRYLVGSAFGTVYKGIWIPDEBNYKIPALIKYLENTSPKANKEILDEAYMAVGVS 779
DB 724 LRKRYLVGSAFGTVYKGIWIPDEBNYKIPALIKYLENTSPKANKEILDEAYMAVGVS 783
QY 780 PYSRLILGICLTSTVQLVTOLMPYGCILLDHVENRGLSGODLLNMCQIAKMSYLEDV 839
DB 784 PYSRLILGICLTSTVQLVTOLMPYGCILLDHVENRGLSGODLLNMCQIAKMSYLEDV 843
QY 840 RLVRHDLAARNVLYKSPNHVKITDGLARLLDIDETEHADGKVPILKMALESILRRF 899
DB 844 RLVRHDLAARNVLYKSPNHVKITDGLARLLDIDETEHADGKVPILKMALESILRRF 903
QY 900 THOSDVSXGVTWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMIWKW 959
DB 904 THOSDVSXGVTWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMIWKW 963
QY 960 MIDSECPREELVSEFSRMAARDPQRFVIONEDLGPASPLDSTFYRSLLEDDMGDLVD 1019
DB 964 MIDSECPREELVSEFSRMAARDPQRFVIONEDLGPASPLDSTFYRSLLEDDMGDLVD 1023
QY 1020 AERTLVVQOQFPCDPRAGAGMWHHRSSSTSSGGDLTLGLEPSEBEPASPLAPSE 1079
DB 1024 AERTLVVQOQFPCDPRAGAGMWHHRSSSTSSGGDLTLGLEPSEBEPASPLAPSE 1083
QY 1080 GAGSDVDGDLGMAAGLGLSLPTHDSPLGRYSEDPVPLPSETDGVAVPLJCSPOEY 1139
DB 1084 GAGSDVDGDLGMAAGLGLSLPTHDSPLGRYSEDPVPLPSETDGVAVPLJCSPOEY 1143
QY 1140 VNOBVPAPPSPREGLPAARPAAGATLERPKTSLPGKNGVVKVAFGAVENTPEYLYP 1199
DB 1144 VNOBVPAPPSPREGLPAARPAAGATLERPKTSLPGKNGVVKVAFGAVENTPEYLYP 1203
QY 1200 OGGAAPQHPAPSPAPADNLYMDQDPPERGAPESTKGTPTANPEYLGDLV 1255
DB 1204 REGTASBPHPAPSPAPADNLYMDQDPPERGAPESTKGTPTANPEYLGDLV 1259

```

```

RESULT 3
Q8COE7 ID Q8COE7 PRELIMINARY; PRT; 881 AA.
AC Q8COE7;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE V-erb-b2 erythroblastic leukemia viral oncogene homolog 2
DE (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK031542; BAC27442.1; -
FT NON TER
SQ SEQUENCE 881 AA; 97501 MW; 5D5042BE9F80836 CRC64;

Query Match 61.9%; Score 4216; DB 11; Length 881;
Best Local Similarity 88.3%; Pred. No. 8.4e-309;
Matches 778; Conservative 39; Mismatches 64; Indels 0; Gaps 0;

QY 375 APLPESFDGPAANTAPLOPEOLQVETLEETGYLYISAMPDSLPLDLSVFNQVIRGR 434
DB 1 APLPESFDGPAANTAPLOPEOLQVETLEETGYLYISAMPESFOLDLSVFNQVIRGR 60
QY 435 ILHNGAYSLTLOGLSIWLGLSLBELSGALIIHNTHLCFVHTVPMDOFLRNPHOALL 494
DB 61 ILHNGAYSLTLOGLSIWLGLSLBELSGALIIHNTHLCFVHTVPMDOFLRNPHOALL 120
QY 495 HTANRPEDECVBGLACHOLCARGHGMGPQCVCNCSQFLRGOECVEBECRYLQGLPREY 554
DB 121 HTANRPEDECVBGLACHOLCARGHGMGPQCVCNCSQFLRGOECVEBECRYLQGLPREY 180
QY 555 VNARHCLPCHPECPONGSVTCFGEADQVCACAHYKDPFCVARCPBGVCKPDLSTYMPWK 614
DB 181 VNARHCLPCHPECPONGSVTCFGEADQVCACAHYKDPFCVARCPBGVCKPDLSTYMPWK 240
QY 615 KPDEBGAOCPCPINCTHSCVDLDDKCPABORASPLTISIAVVGILLVVLGVFGL 674
DB 241 KPDEBGAOCPCPINCTHSCVDLDDKCPABORASPLTISIAVVGILLVVLGVFGL 300
QY 675 IKRROKIRKYMTRLLOTELVEPLTPSGAMPNOAMILKETEILRKYLVGSAFGTV 734
DB 301 IKRROKIRKYMTRLLOTELVEPLTPSGAMPNOAMILKETEILRKYLVGSAFGTV 360
QY 735 YKGIWIPDEBNYKIPALIKYLENTSPKANKEILDEAYMAVGVSPTSRLILGICLTSTV 794
DB 361 YKGIWIPDEBNYKIPALIKYLENTSPKANKEILDEAYMAVGVSPTSRLILGICLTSTV 420
QY 795 QLVTOAMPYGCILLDHVENRGLSGODLLNMCQIAKMSYLEDVRLVRHDLAARNVLYK 854
DB 421 QLVTOAMPYGCILLDHVENRGLSGODLLNMCQIAKMSYLEDVRLVRHDLAARNVLYK 480
QY 855 SPNHVKITDGLARLLDIDETEHADGKVPILKMALESILRRFTHOSDVSXGVTWE 914
DB 481 SPNHVKITDGLARLLDIDETEHADGKVPILKMALESILRRFTHOSDVSXGVTWE 540
QY 915 LMFEGAPYDGIIPAREIPDLLEKGERLPQPICTIDVYMIWKW MIDSECPREELV 974
DB 541 LMFEGAPYDGIIPAREIPDLLEKGERLPQPICTIDVYMIWKW MIDSECPREELV 600
QY 975 EFSHMAADPQRFVIONEDLGPASPLDSTFYRSLLEDDMGDLVDAERTLVVQOQFPCD 1034

```



Db 601 EF8BMAPDPRFVITQNEDELSPSSPMDSTFYRSLLEDDDMGELVDABEYLVPOQGFSPD 660

Qy 1035 PARGAGGVHRRHSSSTRSGGDLTLGLEPSEBEAPRSLAPSEAGSDVFDGLMGGA 1094

Db 661 PALGTSTARRRRHSSSARSGGLTLGLEPSEBEPRSPAPSEAGSDVFDGLMGGA 720

Qy 1095 AKGLQSLPTDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNOQDVRPQSPRE 1154

Db 721 TKGLQSLPHDLSPLOQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNOQDVRPQSPRE 780

Qy 1155 GPLPAPBAGATLRLPRTLSGKNGVYKDVAPFGAIVENPEYLTPOGGAAPQHPPPAFS 1214

Db 781 GPPPLRPAGATLRLPRTLSGKNGVYKDVAPFGAIVENPEYLTPOGGAAPQHPPPAFS 840

Qy 1215 PAFDNLTYMDQDPPERCAPSTFKGTPTAENPEYLTGLDVPY 1255

Db 841 PAFDNLTYMDQNSEQGPSTFKGTPTAENPEYLTGLDVPY 881

RESULT 4

090X70 PRELIMINARY; PRT; 1209 AA.

AC 090X70, (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBlrel. 23, Last annotation update)

DE Epidermal growth factor receptor.

GN EGFR.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN NCBI

RP SEQUENCE FROM N.A.

RC STRAIN=Fishar; TISSUE=Liver;

RX MEDLINE=90258888; PubMed=2342466;

RA Petch L.A., Harris J., Raymond V.W., Blasband A.J., Lee D.C.,

RA Earp H.S.;

RT "A truncated, secreted form of the epidermal growth factor receptor is encoded by an alternatively applied transcript in normal rat tissue.";

RL Mol. Cell. Biol. 10:2973-2982(1990).

RL (2)

RP SEQUENCE FROM N.A.

RC STRAIN=Fishar; TISSUE=Liver;

RA Petch L.A.;

RL Submitted (NOV-1991) to the EMBL/Genbank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Fishar; TISSUE=Liver;

RA Gutteridge K., Dawson T.L., Earp H.S.;

RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.

RX EMBL, M37394, AAP14008.1, -.

DR HSP, P11362, IFGK.

DR InterPro: IPR000494; EGFR\_L domain.

DR InterPro: IPR006211; Furin-like.

DR InterPro: IPR006212; Furin repeat.

DR InterPro: IPR000719; Prot Kinase.

DR InterPro: IPR001245; Tyr\_Kinase.

DR Pfam: PF00757; Furin-like; 1.

DR Pfam: PF00659; Kinase; 1.

DR Pfam: PF01030; Recep\_L domain; 2.

DR PRINTS: PR00109; TYRKINASE.

DR ProDom: PD000001; Prot\_Kinase; 1.

DR SMART: SM00261; FU\_3.

DR SMART: SM00219; TYKIC; 1.

DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.

DR PROSITE: PS00101; PROTEIN\_KINASE\_DOM; 1.

DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.

KW ATP-binding; Kinase; Receptor; Transferrase; Tyrosine-protein kinase.

SEQUENCE 1209 AA; 134891 MW; 96PEB76CC1B773 CR664;

Query Match 46.5%; Score 3172; DB 11; Length 1209;

Best Local Similarity 50.2%; Pred. No. 6,7e-230;

Matches 644; Conservative 166; Mismatches 348; Indels 124; Gaps 27;

Qy 3 LAALCRGILLALPPGA-ASTQVCTGTDKMLRLPASPETHLDMLRLHYOCQVVOGLLE 61

Db 15 LAALCAAG-----GALBEKVCQSTNRRLTQGLTFEDHFLSLQMFNNCEVVLGLE 66

Qy 62 LTYLPTNASLFLQDIOEVGYVLLAHNOVRVPLQRLIVRGTLQFEDNYALAVLQNGD 121

Db 67 ITVQRNYDSLTKTIOEVGYVLLAHNTVERIPLFNLIQIRGNALYENTYALAVLSN-- 124

Qy 122 PLNNTPTVAGSAGGLRELOBSLTILKGGVLIQRNPOLCYODTILMKDIFHQNQLAL 181

Db 125 -----YGTNTGTGLRELMRNLIQETLIGAVRSNNPILCNMTTIQMRDVI--QDVFISN 175

Qy 182 TLIDTNRS-RACHPSMCKSGSRGSSSDCOSLRTVCAGGCA-RCKGPLPTDCHEQ 239

Db 176 MSMDVQRHLTGCKRCDBSCNGSCMGRENCQTLTITCAQCCSRRCRGSPPDCCHNQ 235

Qy 240 CAAGCTGPKASDCLAHFNHSGI CELHCPALVTYNTDTEBSMNPBGRVTGASCVTAC 299

Db 236 CAAGCTGPRSDCLVCHFRFDEATCKDTCPLMLNPTTYQMDVNPBGKYSGCATCVKCK 295

Qy 300 PNYVLTSDVSCVLNQLHNOEYTAEDTQRCCKSCRCARVCYGLMEHLREYRAVTS 359

Db 296 PNYVLTSDVSCVLNQLHNOEYTAEDTQRCCKSCRCARVCYGLMEHLREYRAVTS 354

Qy 360 NIEPAGCKKIFGSLAFLPESFDGDPASNTAPLOPELOVFEETLEETLYLISAMPDL 419

Db 355 NIEPAGCKKIFGSLAFLPESFDGDPASNTAPLOPELOVFEETLEETLYLISAMPDL 414

Qy 420 PDISVFNQIVIRGRLHNGAVSLTLOGLISWGLSLRLSGLALIHNNTHLCFVHT 479

Db 415 TDLHAFENLEIRIGRTKHQGFSLAVAGLNTISGLSLKXISIDGVDIISGNRLCYANT 474

Qy 480 VPMDOLEPRNHOALLHPANPEDECVSEGLACQLCARGHWGPGPQCVCVSCQFLQGE 539

Db 475 INMKKIFGTNQTQKTKIMNNAEBCDCAATNHVCNPLCSBEGCWGEPDVCSCQVSKRE 534

Qy 540 CVEBCRYLQGLPREYVNAARHLCFCHPECOPONGSVTFEGEADOCVCAHYDPPFCVAR 599

Db 535 CVDKCNILBEPREPEFENSCICQHPCLPQTMNITTGSGPNDICAHVVDGPHCVKT 594

Qy 600 CPSEGVKPDLSYMPIMKFPDEEGACQCPINCTHSVDLDKCPAEGRASP-LTISIISAV 658

Db 595 CPSEGVKPDLSYMPIMKFPDEEGACQCPINCTHSVDLDKCPAEGRASP-LTISIISAV 651

Qy 659 VGILLVYVGLVNGI-LIKRQOKIRKTYMRRLQETELVEPLTPSGAMPQAOIRLIKE 717

Db 652 VGILLVYVGLVNGI-LIKRQOKIRKTYMRRLQETELVEPLTPSGAMPQAOIRLIKE 710

Qy 718 TELRKVYVGLSGAGFYTKGIWI PDGENVKIPVAIKYLRNTSPKANKELIDEAYVMAVG 777

Db 711 TELRKVYVGLSGAGFYTKGIWI PDGENVKIPVAIKYLRNTSPKANKELIDEAYVMAVG 770

Qy 778 GSPYVSRLLGICLTSTVQVLTOLMPYGLDHYRENNRGISODLNMCMQIAKMSYLE 837

Db 771 DNHVHVCRLGICLTSTVQVLTOLMPYGLDHYRENNRGISODLNMCMQIAKMSYLE 830

Qy 838 DVRLVHRDLAARVLYKSPNHVKITDPGLARLLDIDTEYHADGKVPKIMMALESILRR 897

Db 831 DVRLVHRDLAARVLYKSPNHVKITDPGLARLLDIDTEYHADGKVPKIMMALESILRR 890

Qy 898 RFTHSDVMSGYTVWELMTFGAKPYDGI PARBIPLDLKGERLPPOPICTIDYVMIMVK 957

Db 891 RFTHSDVMSGYTVWELMTFGAKPYDGI PARBIPLDLKGERLPPOPICTIDYVMIMVK 950

Qy 958 CMWIDSECRPRFELVSEFSRMAPDPRFVIO-NEDLGASPLDSTFYRSLLEDDDKGD 1016

Db 951 CMWIDSECRPRFELVSEFSRMAPDPRFVIO-NEDLGASPLDSTFYRSLLEDDDKGD 1010

Qy 1017 LVDAEELVPOQGFPCDPAAGAGVHHRRSSSTRSGGDLTLGLEPSEBEAPRSLA 1076

Db 1011 LVDAEELVPOQGF-----NSPST-----SRTPLL 1036

```

Qy 1077 PSEBAGSDVFPDGLAMGAAGKIQSLPTHPDSEFLQRYSEDPTVPLPSET--DGYVAPLTCS 1134
Db 1037 S8LSANSN-----S8TVACINRNGSCRVKEDAPLQRYSSDPTSVLTEDNIDITFL----- 1086
Qy 1135 POPEYVNOVDVRRPQPPSPREBGLPAPAPAGATL-----ERPRTLSPGKNGVVKDVFACG 1189
Db 1087 PVEPEYINQ-----SVP-KRPAGSVQNPVYHNQPLHAPAGDLYHQN--PSHN 1130
Qy 1190 AVENPEYL-TEQGAAPQPHPPAPFADNLYYWDQ-----DE-----PERGA 1232
Db 1131 AVSNPEYLTINQ-----PTCLSSGFPDSSALMIQGSNQMNDNEDYQODFPFKAK 1181
Qy 1233 PPSTFKGTPTAENPEYLGIDVP 1254
Db 1182 PNCIFKQ-PTAENAEYLRVAP 1202

RESULT 5
Q8M18 PRELIMINARY; PRT; 1209 AA.
ID 08M18
AC 08M18;
AT 01-OCT-2002 (TReMBLrel. 22. Created)
DT 01-OCT-2002 (TReMBLrel. 22. Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23. Last annotation update)
DS Epidermal growth factor receptor.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxId=9623;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim J.G., Vallet J.L., Nonneman D., Christenson R.K.;
RT "Characterization of uterine epidermal growth factor receptor
RT expression during the oestrous cycle and early pregnancy in pigs.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY117054; AA077472.1; -.
DR InterPro; IPR000345; Cytc_heme_bind.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Set_tmr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_Kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00219; Tyrc; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Receptor; Transferase.
SQ SEQUENCE 1209 AA; 133531 MW; 268B3FB1B36F90F CRC64;

Query March 46.3%; Score 3152.5; DB 6; Length 1209;
Beat Local Similarity 49.7%; Pred.No.2e-228;
Matches 632; Conservative 178; Mismatches 344; Indels 117; Gaps 23;

```

```

Qy 186 TNRBACHPCSPMCKGSRGWSSESDQSLTRTVACAGCA-RCKGPLPTDCHECCAGC 244
Db 181 QSLGSCPKCDPGLCNGSCWAGKEMCKTKRVICAQCSGCRGRSPSDCCHNCCAGC 240
Qy 245 TGRPHSDCLACLNHNSGICELHCPALVTYNTDPEESNPNEGRTFPCASCTAPRYNTL 304
Db 241 TGRPSDCLVCRARRRDEATCKDTCPLMLTYPTTYQMDVNPGLKYSFPAATCYKCPRYNV 300
Qy 305 STDVSGCTLVCPLNQEVTAEDGTORCEKSKPCARVCYGLAMEHLRVRVATVSNIOEF 364
Db 301 VTDHSGCYRACSSDSYER-BEDGYRKCKCGPCGCKVNGSIGISFKDTLSINAINIKGF 359
Qy 365 AGCKKIFGSLAFDESEFDGDPASNTAPLQEPQLOVFETLEITGYLYISAMPDSLPLDSV 424
Db 360 RNTSISGDLHLILVAFRGDSFTRTPLDPKXLDILKTKVEITGFLLIQAMERNRTGLHA 419
Qy 425 FQNLQVIRGLIHNAGVSLTQIGIGISMLGRLRELSSGLALIHNNHLCVHVHPMDQ 484
Db 420 FENLEIRGRKQGFSLAVAGLDIASLGRSLKEISDGVIVSGNNLCYANTISWK 479
Qy 485 LFRNPHQALLHTANRPEDECVEGELACQLCARHGMCPPTQCVNCSQFLRGCEVEEC 544
Db 480 LFGTASQKTKIINRSEKCAMGHICNPLCSSGCGEPERDCMSCNFSKGEVEEC 539
Qy 545 RVLQGLPREYVNAHCLPCHCECPONGSVTCFGEADQVACAHYKDPFCVARNCPSCV 604
Db 540 NVLEGEPRFEYENAEVCQCHPECLPQANKVTCWGRGPDSCVRCAHYIDGPHCVTKCPADI 599
Qy 605 KPDLSYPIWKPPEBEGACQPCPINCNTCHSCVDLDDKGPACQORASPLTISAVY-GILL 663
Db 600 AGENSTL-IWKFADANVCHLCHPNCYVCGVPGLEGCAVDRPKLP-SIANGVGGILL 656
Qy 664 VVVGAVFGLIKRROOKIRKTYMRRLQETELVPLTPSGAMPQAOIRLIKETELRV 723
Db 657 AVVALVGLVLELRK-HIVKRTIRLLQERELVPLTPSEBAPQALIRLIKETEFKV 715
Qy 724 KVLGSAFGTVYKGIWIPDGENVKIPVAIKYLRNTSPKAKEILDEAYVNAVGSFYVS 783
Db 716 KVLGSAFGTVYKGLWIPBEGKVIIPVAIKELRENTSPKAKEILDEAYVNAVSVNPHVC 775
Qy 784 RLIGICLTSTVQVLTQMLPFGCLLDHVRNKGRLGSDLLWCMQIANGMSTLEVRVLVH 843
Db 776 RLIGICLTSTVQVLTQMLPFGCLLDYVREHNDNGSQHLLMWCOIANGMYLDEBRRLVH 835
Qy 844 RDLAARVLYKSPNHYITDPEGLARLLDIDETVHADGKVPKIMMALESTLRRTFTHOS 903
Db 836 RDLAARVLYKTPQHVKITDPEGLAKLGAEXETHABSGKVPKIMMALESTLHRVYTHOS 895
Qy 904 DVMSYGVVWELMTFGAKPYDGIIPARBIPLLLEKGERLPQPPICITIDVYMINVKCMIDS 963
Db 896 DVMSYGVVWELMTFGSKPYDGIIPASISYTLKGERLPQPPICITIDVYMINVKCMIDA 955
Qy 964 ECRFRFELVSESRMARDPQRYVIO-NEPLGRASPIDSTFFYSLSLEDDMGDLVDABE 1022
Db 956 DSRKRFRELLIIEFKMARDPQRYVIOGDESMHLPSPIDSNFYRALMDEEDMEDVVDAD 1015
Qy 1023 YLVPQGFPCDPAPAGAGWVHHNRSSSTSGGDLTLGLEPSEEARPSPLASEGAG 1082
Db 1016 YLVPQGF-HPSTSRTPPLISSLSATST-----PAVACVDRNG-- 1054
Qy 1083 SDVFDGDLGMAAGLSLPTHPDSEFLQRYSEDEPTVPLPSET--DGYVAPLTCSPOPEYV 1140
Db 1055 -----QSYPLKESFLQRYSSDPTGALTESLDITFL-----PAPEYV 1092
Qy 1141 NQPDVRRPQSPREBGLPAPAPAGATL-----ERPRTLSPGKNGVVKDVFACGAVENDE 1195
Db 1093 NQ-----SVP-KRPAGSVQNPVYHNQPLHAPAGDLYHQN--SHSNVAGPDE 1136
Qy 1196 YL-TPQGAAPQPHPPAPFADNLYYWDQ-----DE-----PERGAPSTFK 1238
Db 1137 YLNTFR-----PACINGGLGPAFMAQTGSHQINLNDPVOQAFPPKAKSNGICK 1187
Qy 1239 GTPAENPEYL 1249

```

Db 1188 G-PAENAEYL 1197

## RESULT 6

09EP98 PRELIMINARY; PRT; 1210 AA.  
 ID 09EP98  
 AC 09EP98  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Epidermal growth factor receptor isoform 1.  
 GN Egfr.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/101, 129/SvJ, and 129/SvEvTAC;  
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,  
 RA Sinclair C.S., Pearseall R.S., Green P.J., Yee D., Lampland A.L.,  
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,  
 RA Maibhe N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative Egfr transcripts encoding truncated receptor  
 RT isoforms";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J;  
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,  
 RA Schehl C., Pearseall R.S., Green P.J., Yee D., Lampland A.L.,  
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,  
 RA Maibhe N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative Egfr transcripts encoding truncated receptor  
 RT isoforms";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF275366; AAC28045.1;  
 DR EMBL; AF275364; AAC28045.1; JOINED.  
 DR EMBL; AF275365; AAC28045.1; JOINED.  
 DR EMBL; AF275367; AAC24386.1; --  
 DR HSP; P11362; 1FGK.  
 DR MGD; MGI:95294; Egfr.  
 DR InterPro; IPR000345; CytoC\_heme\_bind.  
 DR InterPro; IPR000494; EGFR\_L\_domain.  
 DR InterPro; IPR006211; Furin-Like.  
 DR InterPro; IPR006212; Furin\_repeat.  
 DR InterPro; IPR00719; Prot\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00757; Furin-like; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 2.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00261; FU; 3.  
 DR SMART; SM00219; TyKc; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C\_1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KM ATP-binding; Kinase; Receptor; Tyrosine-protein kinase.  
 SO SEQUENCE 1210 AA; 134840 MW; 62CD021C9D32B18 CRC64;

Query Match 46.1%; Score 3143; DB 11; Length 1210;  
 Best Local Similarity 49.7%; Pred. No. 1e-227;  
 Matches 635; Conservative 168; Mismatches 354; Indels 120; Gaps 24;

Cy 11 LLLALLPPGAA--STGVCTGTMKRLPASPETHLDMLHYLGQGVQVGNLELYLPTN 68  
 Db 14 LTLTLCAGAGLBERKVCQGTSTNRLTQLGTFFDHLFLSQRMYNCEVVLGNLEITYVGRN 73

Cy 69 ASLSFLDIOEVOGYVLIANHVRQVPLQRIRIVRGTOLEFDNTALAVLNGDPLNNTTP 128  
 Db 74 YDLSFLKTIQEVAGYVLIANTVRIPLNTOIRGNALYENTVALALISN----- 124  
 Cy 129 VTGASPGGLRELQRLSTELLKGGVLLQORNPOLCYOPTILMKDI-----FHKNQNALTLI 184  
 Db 125 -YGTNRGLLELPRNLOEILIGAVRPSNNILCNMDTQWRDYOVNFMNSMSDL--- 180  
 Cy 185 DYNRSRACHPCSPKCKSRGWSSSEDOQSITRTVCAGGCA-RCKGPLPTDCHEQCAAG 243  
 Db 181 -QSHRSPCKPDBPCSPGSCMGSEBNCQKLTIKICAGQCHRCGRGRPSQCHNQCAAG 239  
 Cy 244 CTGPKHSDCLACHFNHSGICELHCPALVYNTDTFBSMPNDEGRYTRGASCVTAQPNY 303  
 Db 240 CTGPRESDCLVQCFQDEATCKDTCPPMLVNPPTTYQMDVNPDEGYSGATCVKCCPNY 299  
 Cy 304 LSTDVSGCTVLCPLHNEVTAEDTQRCCKSRKACVGLGMEHLREVAVTSANI OE 363  
 Db 300 VYTHSGCVRAQCPDYEV--BEDGIRCKCKCDGCRKVCNGIGGEFDLTLSINATNIK 358  
 Cy 364 FAGCKRTFGSLAFPLPESFDCDPAANTAPLOPEQLQVETLEITGYLYISAMPDLPLDS 423  
 Db 359 FKCTAISGDHLILPVAFKDPSFTRTPLEDRLEIKTVAELTGFLLIQAMPDWTDLH 418  
 Cy 424 VFQNLQVIRGRILHNGAVSLTLOGLGISWGLRSLRELGSGLALIHNTHLCFVHTVWD 483  
 Db 419 AFENLEIIRGRTHQGFSLAVVGLANTISLGRSLKEISDQDVIIISGNRNLCYANTIMWK 478  
 Cy 484 QLPENPHQALHTNRPDEDECVGGLACHOU-CARGHCGPPTOCVNSQFLRGEQVEE 543  
 Db 479 KLFTGPNOCKTKIMNRBAKDKAVNHVNCPLSSBGCGPBRDVSQNVSRGECVEK 538  
 Cy 544 CRVLOGLPREVYNARHCLPCHPECOPONGSVTCGPEADQCVACAHYDPPFCVACRPSG 603  
 Db 539 CNILEGPREVENSSEICQHPCCLPQAMNITCGRGDNCICQAHYIDGHCVKTCAG 598  
 Cy 604 VKPDLSTMPYKFPDEBACOPCPINCTHSCVDLDKCKAPAEORASPLTISAVVGLL 663  
 Db 599 IMGENNTL-VKXYDANNVCHLCNANCTYGCAGPGLQCEBWPSPGKIPSTATGVGL 657  
 Cy 664 VYVLGVVFGI-LIRROOKIRKYMRLLOETLEVEPLTPSGAMPNOMQMLIKETELRX 722  
 Db 658 FIVV-VALGIGLFNRHRIYKRTLRRLQRELEVEPLTPSGEAPNHLIKETELRX 716  
 Cy 723 VKVLGSAFGTVYKGIWIPDENYKIPVAIVLRENTSPKANKETILDEAYVMAVGSPYV 782  
 Db 717 IKVLGSAFGTVYKGIWIPSEKIKIIVAIYELAEATSPKANKETILDEAYVMAVDNRYV 776  
 Cy 783 SRLGICLTSTVQLVTOLMPYGLLDHVRNRRGLSGODLLNMCQIAKGSYLEDRVL 842  
 Db 777 CRLLGICLTSTVQLVTOLMPYGLLDYREKDNIGSYLLNMCQIAKGMVLEDRVL 836  
 Cy 843 HRDLAANVLYKSNRYKITDFGLARLLDIDETVHADGCVPIKMMALLESILRRFTHQ 902  
 Db 837 HRDLAANVLYKTHQHYKITDFGLAKLLGAEEKYVHLEGVPIKMMALLESILHLYTHQ 896  
 Cy 903 SDVWSYGVTVWELTFFGAKPYDGI-PAREIPLLEKGERLPPICTIDVYIMVYKMMID 962  
 Db 897 SDVWSYGVTVWELTFFGAKPYDGI-PASDISILKGERLPPICTIDVYIMVYKMMID 956  
 Cy 963 SECPRRRELVSERPMARDPQRFVITQ-NEDLGPASPLDSTFYRSLLEDMDGLVDAE 1021  
 Db 957 ADSRPKRELILERSKARDPQRFVITQ-IGDERMHLPSPTDSNFYALMDEEDMEDVDAD 1016  
 Cy 1022 EYLVPOGCFPCPDPAFPAAGVHHRHSSSTRSGGDLTLGLPEBEERAPSPPLASPCA 1081  
 Db 1017 EYLVPOGCF-----NSPST-----SRPLSSLSA 1042  
 Cy 1082 GSDVFDGLGMAKGLQSLPTHDSPFORSEDPVLPSET--DGVAAPLTCPOPEY 1139  
 Db 1043 TSN-----NSTYACINRNGSCAVKEDAFLOKISSPPTGAVTEDNIDADL-----PVPEY 1092  
 Cy 1140 VNQPDVFPQSPSPREGPLPAARPAQATL-----ERPKTLSFGKGVKVDVAFGAVENP 1194

Db 1093 VNO-----SVP-KRPAGSVONPYTHNOPLHAPGRDLHYON--PHSNAVGNP 1136  
Qy 1195 EYL-TPGOGAARPPHPPAFSPAFDNLTYWDO-----DP-----PERGAPSTP 1237  
Db 1137 EYLTATQ-----PTCLSSGFNSPALWTKGSHQMSLDNPDYQDDFFKTKPKNGIF 1187  
Qy 1238 KGTPTAENPEYLGIDVP 1254  
Db 1188 KG-PTAENAEYLRVAPP 1203  
RESULT 7  
QyYH40 PRELIMINARY; PRT; 1165 AA.  
AC QYH40;  
DT 01-MAY-1999 (TReMBLrel. 10, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
DE Receptor tyrosine kinase proto-oncogene.  
GN XMRK.  
OS Xiphophorus xiphidium.  
OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorphi; Atherinomorpha;  
OC Cyprinodontiformes; Poeciliidae; Xiphophorus.  
OX NCBI\_Taxid=8086;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Rio Purification;  
RX MEDLINE=98241172; PubMed=9582016;  
RA Altschuld J., Scharl M., Winkler C., Wellbrock C., Gomez A., Duesch J.,  
RT "Activation of the Xmrk proto-oncogene of Xiphophorus by  
overexpression and mutational alterations.",  
RL Oncogene 16:1691-1690 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Rio Purification;  
RX Scharl M.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U53471; AAD10500.2; -.  
DR HSP; P11362; IFCG.  
DR InterPro; IPR000345; CYC\_heme\_bind.  
DR InterPro; IPR000494; EGFR\_L\_domain.  
DR InterPro; IPR006212; Purin\_Like.  
DR InterPro; IPR006212; Purin\_repeat.  
DR InterPro; IPR00719; Prot\_Kinase.  
DR InterPro; IPR01245; Tyr\_Kinase.  
DR Pfam; PF00757; Furin-like; 1.  
DR Pfam; PF00659; Kinase; 1.  
DR Pfam; PF01030; Recep\_L\_domain; 2.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_Kinase; 1.  
DR SMART; SM00261; FU; 3.  
DR SMART; SM00219; TYRC; 1.  
DR PROSITE; PS00190; CYCROME\_C; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.  
SQ SEQUENCE 1165 AA; 129614 MW; 7F7E3BD871A74E CRC64;  
Query March 40.5%; Score 2758; DB 13; Length 1165;  
Best Local Similarity 45.9%; Pred. No. 1e-198;  
Matches 584; Conservative 164; Mismatches 384; Indels 140; Gaps 29;

Db 57 VLENLEITTYQENODLSFLOSIGEVGYVLAMNEVSTIPVNLRLRGQVLYEGNFTLL 116  
Qy 116 VLNDGDEPLNNTTPVTGASPGGLRELQSLTEILKGVLIQRNPOLCYQDTILMKDIFHK 175  
Db 117 VMSYQK-NPSSP--DYGVQGLKQLQSLNTEILSGVKSQSHNPLCNVETINMWDIVDK 173  
Qy 176 NQOLATLTLIDTNSRACHPSCPMCKSGRCWGESSEDDQSLRTYCAGGC-ARCGPLPTD 234  
Db 174 TSNPTNMLIHAERQCKDPCGVNCSWAPRGHCQKFTKLLCAQCNNRCRGPXID 233  
Qy 235 CCHROGACGCTGPHSDCLALPHNSGICELHCPALVTYNTDTPFESMPEGRYTFGAS 294  
Db 234 CCHROGACGCTGPRATCLACRDPNDGTCCKDTCPPKIYDIVHQVYDNNIKTFEBA 293  
Qy 295 CYTACPNYLSIDVGSCTLVCPLANOEVTAEEDGTORCEKSKPCARVYGLMEHLEVR 354  
Db 294 CVKRCPSNYVYTB-GACVRSQSAEMLEVD-ENGKRSCKPCGCVKPCVDGIGISLNTI 351  
Qy 355 AVTSANIQERAGCKKIFGSLAFLESFDPGDPASTAPLOPQLOFETLEITGYLYISA 414  
Db 352 AVNSTNGSFNCTKINGDIILNNSFEGDPHYKIGPMDPEHLNLTLYKEITGYLYIWM 411  
Qy 415 MPDLSPLSVFQNLQVTRGRILHNGAYS-LTIQSLGISWLGSLRSLRELSGALILHNTH 473  
Db 412 WPEMNTSLVFOMLIIRKTTTSRGSFVYVQVSHLOMGLSLKEVSAGNVILAKTPQ 471  
Qy 474 LCFVHTVPMDLPANPHQALLHTANRPEDCEVSGELACHOLCARGHCMGPPPTQVNCQ 533  
Db 472 LRYASTINMRRLPRSEDSIEYDART-----ENQCNNESDGCWGPPPTMCVSLH 524  
Qy 534 PLRGQCEVERCRLVGLPREYVNAARCLPCHPECPONGSTTCGPEADQCAACHYADP 593  
Db 525 VDRGRCVASCNLLQGPBREAOVDRCVQGOECLVTQDSTLCYGPBPANCSKCHAFDDG 584  
Qy 594 PFCVARGSPGVKPDLSMPIMKPPDEBGAOCPCINCHSCVDDDDKCPAQRASPLTS 653  
Db 585 PQICRCPHMLGGDTL-IMKTADKMGQCCPCQNCQTCGSCGGLSCRGD-IVSHSL 642  
Qy 654 IISAVGILLVVVGVVFGILIKRQOKIRKYTRRLQETELVEPLTPSGAMPNQAMR 713  
Db 643 AVGLVSLILTVIALLIVLLRRRRRK-RRTIRLLQEBELVEPLTPSGAPRQAFLR 701  
Qy 714 ILKTELARKYKVLGSGAFGVYKGIWIPDGNVKKPAVAKYLRNTPSKANKEIIDEAYV 773  
Db 702 ILKTEFEKQKRVLGSGAFGVYKGIWIPDGNVKKPAVAKYLRNTPSKANKEIIDEAYV 761  
Qy 774 MAGVSPYVRLLGICITSTVQVLTOLMPYCCILDHRENGRLGSDQLMMQMIAGM 833  
Db 762 MASVDHPCVRLGICITSAVQVLTOLMPYCCILDYRQHOERTCGQMLMNCVQIAGM 821  
Qy 834 SYLEDVRLVHRDLAARVAVKSPNHVKITDGLARLLDIDETEHADGKVPIMKMALES 893  
Db 822 NYLEBRHLVHRDLAARVAVKSPNHVKITDGLARLLDIDETEHADGKVPIMKMALES 881  
Qy 894 ILRRRTHQSDVSYGVYTWELMTFGAKPYDGIAPARIPDLLEKGERLPOPPICITIDYM 953  
Db 882 ILQWYTHQSDVSYGVYTWELMTFGSKPYDGIAPAKETASYLENGERLPOPPICITIEYVM 941  
Qy 954 IMVCKMIDSECRPRELVESEFARMADPQRFVIONEDIGPASPPLSTFRSLDEDD 1013  
Db 942 IILCKMIDPESRRFRLVGEFSQMAKDPERSYLVITQ--NLPSRRLRLSLSSDD 998  
Qy 1014 MGDVLDAEYLVPOCGFFCPDPAAGAGMVHRRASSTRSGGDLTLGLBSEBEAPRS 1073  
Db 999 --DIVVDADEYLL-----FKRIN-RGS-----E 1019  
Qy 1074 PLAPSEAGSDVFDGLGMAKAGLSLPTHDPEPLQRYSDPYV-PLPSETDGYAPLT 1132  
Db 1020 PCIPPNGH-----PYRENSIARLYISDPTONALEKIDG----- 1054  
Qy 1133 CSPPPEYVNGPDVAPQ-----PSPRE-----GPLP-AARPAATLBRPKTSLPGKNG 1179  
Db 1055 -----EYVNOGSETSSRLSDIYNPNYEDLTDGNGPVSLSQEAETNFSRPEYLTNONS 1109

```

QY 1180 VKDVFAGGAVENPEYLTCGGAAAPRHRPAPSPAFDNLVYWDOPRERGAPSTFKG 1239
D 1110 L-----PLVSSGSMDDPDY-----QAG-----YQAF-----LPQGLALTGGMF 1144
QY 1240 TPTAENPEYIGL 1251
D 1145 LPAENLEYLGL 1156

RESULT 8
QW6F6 PRELIMINARY; PRT; 1137 AA.
ID QW6F6
AC QW6F6;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Receptor tyrosine kinase (fragment).
GN ERBB4.
OS Gallus gallus (Chicken).
OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hindbrain;
RX MEDLINE=99263203; PubMed=10328884;
RA Dixon M., Lumsden A.,
RT "Distribution of neuregulin-1 (nrg1) and erbB4 transcripts in
RL embryonic chick hindbrain.",
RL Mol. Cell. Neurosci. 13:237-258 (1999).
DR HSSP: P11362; IFGK.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR006211; Furin-like.
DR InterPro: IPR006212; Furin repeat.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR001368; TNFR_c6.
DR InterPro: IPR001451; Tyr_kinase.
DR InterPro: IPR004019; YLP_motif.
DR Pfam: PF00757; Furin-like.
DR Pfam: PF00069; pkinase.
DR Pfam: PF01030; Recep_L_domain.
DR Pfam: PF02757; YLP_2.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Prot_kinase.
DR SMART: SM00261; FU; 3.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00652; TNFR_NGFR_1; 1.
KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
FT NON TER
SQ SEQUENCE 1137 AA; 127927 MW; 4D616436F87D84F CRC64;

Query Match 40.1%; Score 2735.5; DB 13; Length 1137;
Best Local Similarity 47.5%; Pred. No. 4.9e-197;
Matches 542; Conservative 172; Mismatches 345; Indels 83; Gaps 25;

```

```

QY 340 RVCYGLMEHLAEVRAVTSANIOEFAGCKKI FGSIAFLPESFDGDPASNTAPLOPQOV 399
D 180 KACDGIQTSGLVSAQVATSSSNDIKF INCTKINGLILVYGHGDPHTIAAIPEKLN 239
QY 400 FETLEETITGLYISANPDSLPLDSVFONLOVYIGRIILNAGAYSLTLOGLGISWGLASLR 459
D 240 FQTVREITGYLNIQSWPENMTDFRVSNLVTIGRALYSGSLILKQOGLTQFQSLK 299
QY 460 ELGSGALAIHNTLFCVHTVPMDLFRNPHQALHTANRPEDECEGGLACHOLARGH 519
D 300 QISAGNIYITDNLNLCTHTVNTSLFSTPSQKTVIHNRKKAENCTADGAVCNLCSGDS 359
QY 520 CWBPPTQVNCGQPLRQGEVBECEVYLOGLPREYVNAHCLPCHPEQOP-QNGSVTCFG 578
D 360 CWBPPTQVNCGQPLRQGEVBECEVYLOGLPREYVNAHCLPCHPEQOP-QNGSVTCFG 578
QY 579 PEAQVACAHYDPPPCVACRSPGVKQVPLSYMPIKFPPEBAGCOPCINCHSAYDLD 638
D 420 PGPDCCKCFHPDPCVACRSPGVKQVPLSYMPIKFPPEBAGCOPCINCHSAYDLD 638
QY 639 DKCSPAEQ-----RASPL-----TSISAVV--GILLVVLGVVFGILIKRQOKIR 683
D 474 -RG-PASHDCIYYPTWQSTLPQHARTPLIAGYIGLFTIIVMGLTFAYVVRKSKIK-K 530
QY 684 KYTMRLLQETELVEPLTPSGAMPNQAMRIKTELKRYKVLGSGAFGVYKGIWIPDG 743
D 531 KRALRFL-ETELVEPLTPSGAMPNQAMRIKTELKRYKVLGSGAFGVYKGIWIPDG 589
QY 744 ENKIPYALIVALENTSPKANKSILDEAVYMAVGSPYVSRILGICLTSTVQVLTQMLPY 803
D 590 EYKIPYALIVALENTSPKANKSILDEAVYMAVGSPYVSRILGICLTSTVQVLTQMLPY 803
QY 804 GCLLDHYENRGRIGSODLNMCMQAKGMSYLEVRLVHRDLAARVLYKSPNHVKITD 863
D 650 GCLLDHYENRGRIGSODLNMCMQAKGMSYLEVRLVHRDLAARVLYKSPNHVKITD 863
QY 710 FGLARLLDETEYHADGKVPILKMMALISILRRRFTHQSDVMSYGVTWELMTFGAKPY 923
D 924 DGIPARIEPLLEKGERLPQPICTIDVYMIKMMIDSECRFRFELVSEFSRMADP 983
QY 770 DGIPARIEPLLEKGERLPQPICTIDVYMIKMMIDSECRFRFELVSEFSRMADP 829
QY 984 QRFVIONED-LGPASPLDSTFYSRLLEDMDGLVAAEYLVPOQGFPCDPAPAGAGM 1042
D 830 QRFVIONED-LGPASPLDSTFYSRLLEDMDGLVAAEYLVPOQGFPCDPAPAGAGM 1042
QY 1043 VHRHRSSSTRSGGDLTLGLSPSEEARPS--PLAF--SEGAGSDVFDGDLGMAKGLQ 1099
D 889 IDSRNRFVYRDGVAALAEQGV-PWYRAPGCIIEAVVAGATAEIFEDTCNGTLRKQV 947
QY 1100 SLPTHDPSPLOKRSBPTVPLPS-----ETDGYVAPLTCSPQPEYVNOVDVAPQPSF 1152
D 948 ATLAKEDESSQVRSADPTVPIPERVIRGELDEDEQVMTPMKPTDILNPEENPFVSRR 1007
QY 1153 REGPLPARAPAGATLSPK-TLSPGKGVVADV-----ARGAVENPEYLTPQCGA 1203
D 1008 KNGDLQ-----VDPEYHANAENGQPKADEYVNEPLYNTANTLENAYL----- 1054
QY 1204 APOHPPEAPSPAFDNLVYWDOPRERGA--PSPFGKPTP-----AENPE 1247
D 1055 --KNMLPEKAKKAFNDNDYNNHSLPRPSTIQHPYLDGYSTKYFYKQNGRIRPIVAENPE 1112
QY 1248 YL 1249
D 1113 YL 1114

RESULT 9
P79754 PRELIMINARY; PRT; 1328 AA.
AC P79754;

```

DT 01-MAY-1997 (TReMBLrel. 03, Created)  
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
 DE Erib3.  
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;  
 OC Tetraodontidae; Tetraodontidae; Takifugu.  
 NCBI\_Taxid=31033;  
 RN [1]  
 RP SEQUENCE FROM N. A.  
 RX MEDLINE=9177347; PubMed=10077531;  
 RA Gellner K., Brenner S.  
 RT "Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu  
 rubripes."  
 RL Genome Res. 9:251-258(1999).  
 DR EMBL; AF056116; AAC34391.1; -.  
 DR HSSP; P11362; IFGK.  
 DR InterPro; IPR000494; EGFR\_L\_domain.  
 DR InterPro; IPR006211; Furin-like.  
 DR InterPro; IPR006212; Furin repeat.  
 DR InterPro; IPR000719; Prot\_Kinase.  
 DR InterPro; IPR001245; Tyr\_Kinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 2.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00261; FU; 3.  
 DR SMART; SM00219; TyKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 KM ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.  
 SQ SEQUENCE 1328 AA; 148613 MW; A33303258B647B9 CRC64;  
 Query Match 34.1%; Score 2326; DB 13; Length 1328;  
 Best Local Similarity 40.7%; Pred. No. 4.5e-166;  
 Matches 524; Conservative 154; Mismatches 418; Indels 190; Gaps 31;  
 QY 9 WGLLALLLP--GAASV---VCTGDMKRLRSPETHLDMLHLYOGGVVQVQNIEL 62  
 DB 4 WRLILMCVARSRLRAASQTOBAVCPGTONGISTGSGENOVNLMNRKKEIIMGNIEI 63  
 QY 63 TYLPFNASLSPLODIOEVGVYVLAHQVROVPLORLIRVGTOLFEDNVALLVLDNDP 122  
 DB 64 TQISNMFSPFLKTRVETGVYLAHNFPOEIPGLQALVIRGNSLYERRFALSYFLN--- 120  
 QY 123 LNNTPVTGASPGGIRELQRLSLTEILKGVLIQRNPOLCYODTILMKDIFKNNQLALT 182  
 DB 121 ----YFDG--PSGLNQLGMLNLTETILLDGVQIINNKTLRGPWWYWDII-RNNDAIE 173  
 QY 183 LIDNRSRACPCSPCKGSRGCBSSSEDCSLRTVCAAGC-ARCKPRLTDCCHSCCA 241  
 DB 174 IQFNGRGVCH---KSC-GNYCWGPGKDCQILTKTVCAPOCNDRCFSTSPDCHIECA 229  
 QY 242 AGCTGPKHSDCLACLPFHSGICELCPALVTYNTDPESMPNPREGRYTFGASCTYACPY 301  
 DB 230 AGCGKPLDTDFACRLFDSDGACVPCQOTLIYNKQITOMETNPAKYQSGISICVSCPT 289  
 QY 302 NYLSTDVOSCTLVPLHNOETABDGTOR-CEKSKPCARCYGIGMEHLREVAVTSAN 360  
 DB 290 HFV-VDSSCVSVCPDGMKV--ERGSQROCELGGLCPKYCEGTGAL---QROTVDSN 343  
 QY 361 IQBPAGCKKIRGSLAFLPESPDGASNTAPLOPQLOVFTLEITGVLYISAMPDLP 420  
 DB 344 IDSFINKTKIGSGHFLVTGLGDGPKVVPPLADAKKEVFRVREITDIILNIOGWPKLN 403  
 QY 421 DLSPFONQVTRGRIHNGAVSLTLQGISWLGRSLREISGSLALHNHTHLCFVHTV 480  
 DB 404 DLSVPBSLTITQKSLFRFSLMVRITPLTSLGSRSLREISDGSVYISQNHLCYHTIV 463  
 QY 481 PWDQLFRNPH-QALHTANRPEDECVGBGLACHQLCARGHGCPGPTQCVNCSQPLRGQ 539

DB 464 NMQLFGRGSRVANSLSNRPMAECVADGRVCDPLCSGSGMGPGPDCLSCRNYSRGT 523  
 QY 540 CVEECRYLQGLPREYVNAH-CLCPHECQPOGNSVTCFGEADQCVAAHYKDPFCVA 598  
 DB 524 CVAGCHNSGIPREFALNVCVACHCECKQGTAKSTGAGDECACTFRDGPYCMS 583  
 QY 599 RCBSGVKPLSYNPIKFPDEBEGACQCPINCHSCYVDLDKGPACORASPLTISAV 658  
 DB 584 SCPAGVA-DEKGLIFKFPNRBCHBCPCHONCTGCGSGPGINDC---LEARNLTSSGOI 639  
 QY 659 VGILLVVLVVF-----GLIKRQOKIRKXTMTMRLLQETLVEPLTSGAMPNQAQ 711  
 DB 640 TGLAGVPAGLIFCLVFLFGMLYHRGLAIRKRAMRYLESGSFELPQ-GEKGTIVH 698  
 QY 712 MRILKETELRKVKVLSGAGFVYKGIWIPDGENVKIPVALKVLRENTSPRANKELDEA 771  
 DB 699 ARILKPSDLAKIKPLSGGVGTYSKGFWIPBGEIVKLPVALKTIQDSSGRQTFETDHL 758  
 QY 772 YNAGVSPYVSRLLGICLTSTVQVLTQMPYGLLDHVRNRRGLSGQDILLNMCQIAK 831  
 DB 759 LSMGSLDHPYIVRLGICPGTCLQVLTQLSHSGILHEIRQKTSLDPQRLNMCVQIAK 818  
 QY 832 GMSYLEVRLVHRDLAARNLVKSPNNVKITDPEGLARLDIDETBYHADGKVPITKMAL 891  
 DB 819 GMYLIEHRVYVHKVLARNLLKNQDYOVIDYVADLLYPDCKKYVSETKTPIKMAL 878  
 QY 892 ESILRRFTQSDVWSGYVWBLMTFGAKYPDGIPIAREIPDLLEKGERLPQPICTIDV 951  
 DB 879 ESILFRYTHQSDVWSGYVWBLMTFGAKYPDGIPIAREIPDLLEKGERLPQPICTIDV 938  
 QY 952 YMIWKCMTIDSECRPRELIVSEFSHWARDPQFVYVIONEDLPASPLDSTFYRSLLED 1011  
 DB 939 YMWVVKCMIDENIRLPTFKELASDFTMARDPRLYLIVRMG-----ED 982  
 QY 1012 DDMGDVDAERYLVPOGFCPPDPAPAGAGVHHRRHSSSRSGGDLTLGLEPSEENP 1071  
 DB 983 SGMESEFL-----RGSRR-----GLLEADLEDEDEE-- 1008  
 QY 1072 RSPFLAPSEGASDVFDGLGNG--AAKGLQSLPTHDPSPQO-----RYSEDPTV 1118  
 DB 1009 -----GLGDRFATPSLOPSFSWISPCQINSYMTQLRYD----- 1044  
 QY 1119 PLPSETDGYVAPLTCSPQ-ERYNQ-----PVRPQPSPREGL--PAAR 1161  
 DB 1045 -FVNSGGHGYLPMSPSPVDITRQLWYQRELSVRLTPRSAFRRSREALECEDGAQ 1103  
 QY 1162 PAGATLERPKLSFGKGVKQVADVAFGGAIVENPEYLLTPOGGAARQPPHPPAFSPFDLY 1221  
 DB 1104 CAGIFRYR-----FGSERGN-----FOGG----- 1122  
 QY 1222 YMDQPPERGAPESTFKGTPTAENPE 1247  
 DB 1123 --QQRKLSYASSPSSFTMAADEDE 1146  
 RESULT 10  
 Q8AM81 PRELIMINARY; PRT; 1305 AA.  
 AC Q8AM81;  
 DT 01-MAR-2003 (TReMBLrel. 23, Created)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
 DE S1:dz150112.1 (Novel protein similar to Erib3 (v-erb-b erythroblastic leukemia viral oncogene homolog, neuro/glioblastoma derived oncogene homolog)).  
 GN S1:dz150112.1.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
 OC Cyprinidae; Danio.  
 NCBI\_Taxid=7955;  
 RN [1]







```

QY 26 CTGTDMLKRLPASPETHLDMKRLH YGCGVVOGULETLYLPTNASLSFLDIOGVGYVL 85
DB 1 CIGTNGMSVYANREYKMLRDRYTCYVDGNLEITWIONITDNLQHIRRETVGYVL 60
QY 86 IAHNOVQOVLORIRIYRGTOF-----EDNYALAVLNDGPNLNTTPTVTCASGGGLREL 140
DB 61 ISLYDLPOVILPRLOIRGRITTYKLNKWEBAVGLFV-----SFSHMTL 104
QY 141 QLASLEILKGLVLIORNPOLCYODTILMKDI -FKHKNQALATLIDTNRSHACPCSPMC 199
DB 105 ELPLRLRILOGSVFPNNYNLCHMKSIWBEIILAPQSMGYTNFSPREHCPCHSC 164
QY 200 KGSRCWGESSEDCSLTRTVACGCA--RCGPLYPTDCCHQCAAGCTGPHSCLACLH 257
DB 165 EVG-CMGECAHNCORFSLNCSPOCSQGRCPKPRECHLFCAGGCTGPTOSDCLACKN 223
QY 258 FMSGIELNCPALVYNTDTFESMPNEGRYTCASCTVTCPTNYLSTDVSGCTLVCP 317
DB 224 FYDDGVKQECPPMOIYNPTNYFWEPNPDGKAYGATVRCRP-EHLKDNAGACVRCRCPK 282
QY 318 HNGEVLADGTORCEKSKPCARVCYGLGMEHLREAVTSANIQEFAGCKKIFGSLAFL 377
DB 283 GKRPQNE-----CVPCKGVCPKTCRGGIYH-----SDNIGNYKDCITLIESGLEIL 329
QY 378 PESFDGDPASNT-----APLOPELOVEETLEITGYLYISAMPDLSPLDSVFORLQ 429
DB 330 DQSFDFGQOYVYTNPSFGRYIKIDPRLEVPSTYKEITGFINIQAHNFTLWYFRLE 389
QY 430 VIRGRIILHNGAY-SLTLQGLGISWLGLSLRELSGLALIHNNHLCVHTVPNDQLFN 488
DB 390 VVGGRQKLENFASVYIVKTSLSKSLRLKRVNVSGLVIIENSLCCEVEDIDMSSEIKS 449
QY 489 PHQALHTANRBECEGEGLAGHOLCARHGMGPGPTQCVNCSQPLRSGOESVECRVLQ 548
DB 450 SDHEVMOYKNNATTECHBEHGECSQKAGCMGKPPQCECKNVKTKGCLDSCK--- 506
QY 549 GLPREY-VNARHCLPCHPECOPONGSVTCFGRPADQCAVACHYDPPFCVARCP----- 601
DB 507 SLRPLYSVSKTCDQHQEKD-----FCYGPNEIDNCSQNNVQDGRCAVACEPTTKHAM 561
QY 602 -----SGVXPULSTYMPKFPD----- 618
DB 562 NGTCINCHKTGCGRPRDTIAPDGCISDKAIIIGSDAKIERCLMKOESCEPDGYSDYVL 621
QY 619 -EEG----- 621
DB 622 QEBEPLKOLSGKAVCRKCHPRCKCTGYFHBOFCQCTGYKKGQCEDECPDFYANEE 681
QY 622 -ACQPCPINCT-----HSCVDL-----DD-----KGCBAEQ----- 646
DB 682 TRICLPCHQBCRGCHGLDHDHCECNLKLFBGDPYDNTATTTCVNSCNASHPRYKFRPOEA 741
QY 647 -----RASPLTJISAVVGLLVVLGVFGLI-----LIKRRQOKIRKXTM 687
DB 742 GKIGPYCSADSMOGLRIEPTQYVCIWGSVMAILLICVVGIAFLVPSRHKNKDKADAVKM 801
QY 688 RLLQETELVEPLRPSGAMPNOAQRILKTELKRVKYLGSARGTYVYKGIWIPGEVVK 747
DB 802 TMAAGCEDSEPLRPSNVGPNLTKRIITKEABIRGVLGAGAFRVRKGVWMBEGSEVVK 861
QY 748 IPVALIKVLENTSPKANKEILDEAVYMAVGSPYVRLGICTSTVOLVTOLMPEYGLL 807
DB 862 IPVALIKVLENTSPKANKEILDEAVYMAVGSPYVRLGICTSTVOLVTOLMPEYGLL 921
QY 808 DHVENRGLSSODILLNWCQOIAKMSYLEVRLVHRLAARVLYKSPNNHKTDPGIA 867
DB 922 DYVANNKOKISSKALLNWSIOIARGMAVLEERLVRHRLAARVLYVQPPSCVKTIVFGIA 981
QY 868 RLLIDIDREYADGQVRIKXMALESILRRPTTHQSDVWSGVYTWELMTGAKXYDIP 927
DB 982 KLLPDSBYRPAAGGKMPIKMLALECIRHRRFTSKSDVAFGITTWELLTGARYEYENP 1041
QY 928 AREIPDLLEKGBRLPQPICTIDVYMWKCMIMSEGRPRFRELVSFSEWADPQRFV 987

```

```

DB 1042 AKQVPELIEIGHKLPQPDICSLDYCILLSCWVIDADARPFKOLAETFAKARDPGRYL 1101
QY 988 VIGNEDGPASPLDSTFYRSLIEDDMGDLV----- 1018
DB 1102 MI-----PQDKFRRLPSYTNQDKDLIRTLAPVAMAAAAAAGASNVDPSTIA 1152
QY 1019 DABEYLVQCGFFCPDPAFAGAGVHHRRSSSTRSGGDLTLGLEPSEEAPRS----- 1073
DB 1153 ETDREYLOPKTRPSIMLPGSA-----VEPS-DEMPKSLRYCK 1188
QY 1074 -PLAF-----SEGAGDYVDGDLGMAAGLQSLPTHDSPLORYSEDPVPLPSETDGYVA 1129
DB 1189 DPLKPDDETGHGKEV-----GVGGIR-----LNLPLDEDDYLM 1222
QY 1130 PLTCSPOEYVNOQDVARPPSPREGLPLPARPAGATLERKTLSPGKNGVAVDFARFG 1189
DB 1223 P-TCQSQ-----NQS-----TPG-----YMDLIGVPA 1243
QY 1190 AVENPEYL-----TPQGAAPOPHPAPFAPFAPDNLYYMDODPERGAPSTFKGT 1240
DB 1244 SYDNPEYLMGSTOALAGLQAGSMG-PHTPP-----PPTPTPGM 1280
QY 1241 PTAENPS 1247
DB 1281 PTHQSQ 1287

RESULT 12
Q8BMLMO ID Q8BMLMO PRELIMINARY; PRT; 1377 AA.
AC Q8BMLMO 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE CG10079-PA.
GN EGFR OR CG10079.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amentadides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Baeu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Bokora D., Botchan M.R., Bouck J., Brockstein P., Broctier P.,
RA Buttle K.C., Buzam D.A., Butler H., Cadieu B., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Duthin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Fowler C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glaeser K.,
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,
RA Jaimel M., Kalluh F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,
RA Kimmel B.B., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Patel J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

```

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,  
 RA Svirskas R., Tector C., Turner R., Venter G., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Zibbe R.A., Myers B.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*,"  
 Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banazon J., An H., Baldwin D., Banazon J., Beeson K.Y., Blaum D.A.,  
 RA Carlson J.W., Center A., Chapple M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Doree V., Doup L.E., Doyle C., Dresnek D., Fattan D.,  
 RA Ferrera S., Frise E., Galle R.F., Gary N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoekins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nuno J.,  
 RA Pacib J., Pargass V., Park S., Patel S., Pfeiffer B., Nuno J.,  
 RA Prounhevong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of *Drosophila melanogaster* genome,"  
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield B.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RT "Annotation of *Drosophila melanogaster* genome,"  
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB003454; AAM70919.1; -  
 DR FLYBase: FBgn003731; Egfr.  
 DR InterPro: IPR000345; CyC\_heme\_bind.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR006211; Purin\_Like.  
 DR InterPro: IPR006212; Purin\_repeat.  
 DR InterPro: IPR00719; Prot\_kinase.  
 DR InterPro: IPR002290; Ser\_thr\_kinase.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF00069; Kinase; 1.  
 DR Pfam: PF01030; Recep\_L\_domain; 2.  
 DR PRINTS: PR0109; TYRKINASE.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00261; FU; 7.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR SMART: SM00219; TyKc; 1.  
 DR PROSITE: PS00190; CYTOCHROME\_C; 3.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00101; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW ATP-binding; Transferase.  
 SQ SEQUENCE 1377 AA; 153601 MW; 38AFAAAB9D5C954 CRC64;

Query Match 29.5%; Score 2013.5; DB 5; Length 1377;  
 Best Local Similarity 33.1%; Pred. No. 1.7e-142;

| Matches | 481; | Conservative  | 185;  | Mismatches                | 435; | Indels | 353; | Gaps | 42; |
|---------|------|---|-------|---------------------------|------|--------|------|------|-----|
| QY      | 9    | WGILLALPPGAAS   | ----- | QVGTDMKRLPASPETHLDMLRYOGC | 53   |        |      |      |     |
| DB      | 21   | WVLLILLACASITSSSVSNAGVNDGNMKNVCIGTSSRLSVSPKHHYRNLAIDRYTNC   | ----- | -----                     | 80   |        |      |      |     |
| QY      | 54   | QVQVGNLELYLPT-NASLSFLDIDQVGVITLANQVROVPLRLRYRGTOLEF----     | ----- | -----                     | 108  |        |      |      |     |
| DB      | 81   | TYVDGNLELTLWLPENMLDLSFLNIREVTSYIILSHVDVKKVPEKLIIRGRTLFSLV   | ----- | -----                     | 140  |        |      |      |     |
| QY      | 109  | -EDNVALAVLNDNDPLNNTTPVVGASRGSLRLQSLTILNGSVLIQRNPOLCYODTI    | ----- | -----                     | 167  |        |      |      |     |
| DB      | 141  | EEKVALFV-----   | ----- | -----                     | 184  |        |      |      |     |
| QY      | 168  | LWMDIFKNNQALTLIDTRSRACHPCSPWCKSGRCMGESSEDCQSLTRTVACAGCA--   | ----- | -----                     | 225  |        |      |      |     |
| DB      | 185  | QNSEIYSNGDAYNDFTPARECEPCCHSCHTG-CWGGPKKQCFKSLTCSPOCAGG      | ----- | -----                     | 243  |        |      |      |     |
| QY      | 226  | RCKGPLETDCCHQCAAGCTGPRHSDCLAFHNSGICELACPALVTYNTDTESMNP      | ----- | -----                     | 285  |        |      |      |     |
| DB      | 244  | RCYGPXKRECCHPFCAGGCTGPTQKCIACKNFDEBVCBECPMKRYPTTVLETNP      | ----- | -----                     | 303  |        |      |      |     |
| QY      | 286  | EGRYTGASCTVACPNVYLTSDVSGCTLVCPLENNQVTAIDGTORCEKSKRCARVGYL   | ----- | -----                     | 345  |        |      |      |     |
| DB      | 304  | EGRYATGATVKECP-GHLIRDNACVRS-CPQDRMDGCB-----                 | ----- | -----                     | 357  |        |      |      |     |
| QY      | 346  | GMEHLEVRVAVTSANIOEFAGCKIFGSLAFLESPDQ--DPASNTA-----          | ----- | -----                     | 397  |        |      |      |     |
| DB      | 358  | TVLH-----   | ----- | -----                     | 409  |        |      |      |     |
| QY      | 398  | QVETLEETITGYIYISAMPDLSVDFONLQVIRGILHNGAV-SLTQGLGISWGLR      | ----- | -----                     | 456  |        |      |      |     |
| DB      | 410  | EVSSTVEITGYINIEGTHQPNRLSYFRNLETIHGQLESMSPALAIVSSYSLEMR      | ----- | -----                     | 469  |        |      |      |     |
| QY      | 457  | SLREISGLALHHNTHLCVHTVPMQDLFRNPHQALHTANREDECVGSLAQHQA        | ----- | -----                     | 516  |        |      |      |     |
| DB      | 470  | NLQISSGVVIGNRDLCTVSNTRMPATQCEPQKWNENLRADLCENKGTICSDGN       | ----- | -----                     | 529  |        |      |      |     |
| QY      | 517  | RGHCWGPPTQCVNCSQFLRGOECVEECRYLQGLPREYVANAHLCPHPECOPNGSVTC   | ----- | -----                     | 576  |        |      |      |     |
| DB      | 530  | EDCGWAGTQCLCTKNFNGTCLADCGYISNAK--FDRRTCKICHPER-----         | ----- | -----                     | 581  |        |      |      |     |
| QY      | 577  | FGPADQCAAHKYPDFCVARCP-----                                  | ----- | -----                     | 605  |        |      |      |     |
| DB      | 582  | NGAGADHCQCVVRDQHCVCSECPKKNYDRGVCRECHATCDCTGPKXTIGIACATTC    | ----- | -----                     | 641  |        |      |      |     |
| QY      | 606  | -----   | ----- | -----                     | 628  |        |      |      |     |
| DB      | 642  | NLAINNDATVKKCLLKDCKPD-GY--FWBYVHPQSGSLKPLAGRAVCRKCHPLCELC   | ----- | -----                     | 698  |        |      |      |     |
| QY      | 629  | -----   | ----- | -----                     | 634  |        |      |      |     |
| DB      | 699  | TNYGYHQCYSKTHYRRQRCETECPADHYTDEQRBCPCQHPBCNCTGPGADDCSC      | ----- | -----                     | 758  |        |      |      |     |
| QY      | 635  | -----   | ----- | -----                     | 653  |        |      |      |     |
| DB      | 759  | RNFPLFANETGPRVNSTWENCTSKCPLBEMHVVNYQYATAGPYCAASPPRSKITTANDV | ----- | -----                     | 818  |        |      |      |     |
| QY      | 654  | -----   | ----- | -----                     | 707  |        |      |      |     |
| DB      | 819  | NMIFITIGAVLVPITCILCV--TYICRQKQKAKKEVTKMTALSGCDSSEPLRPSNIGA  | ----- | -----                     | 876  |        |      |      |     |
| QY      | 708  | NQAGRIKETELKAVLVGSGAGFVYKGIWIDGENVKIPAIKVLRENTSPYANKKI      | ----- | -----                     | 767  |        |      |      |     |
| DB      | 877  | NLCRLIVDAELARKGVLMGAFGRYKGVYEGEVNKKPAVKIKELLSKTAGSSSEF      | ----- | -----                     | 936  |        |      |      |     |
| QY      | 768  | LDEAYVAVGSSPYVSLIGITLSTVQVLTQMPYCGLLDVBENRGLASQDILLNCM      | ----- | -----                     | 827  |        |      |      |     |
| DB      | 937  | LREAYIVASVEHVAILKILAVCMSSQMLITQLMPLCLDLYVNNRDKISSKALLNST    | ----- | -----                     | 996  |        |      |      |     |
| QY      | 828  | QIAGKMSYLEVRLVYHDDLAABRNLYKSPNIVKLTIDEGIARLIDIDETEVADGKVIK  | ----- | -----                     | 887  |        |      |      |     |
| DB      | 997  | QIAGKMSYLBKRLVHDDLAAARNLVOTPSLVKLTIDEGIARLIDIDETEVADGKVIK   | ----- | -----                     | 1056 |        |      |      |     |

QY 888 WMALESILRRRTTQSDVMSYGVYTWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPPIC 947  
 DB 1057 WLAECIRNRVFTSKSDVMAFGVITWELTTFQGRPHENIPANDIPDLIEVGLKEQPEIC 1116  
 QY 948 TIIDYIMVWKWMDSECRPRELVESEFSRMAHDPQFVYIQNDLGG--PASPLDSTFY 1005  
 DB 1117 SLDIYCTLLSCMHLDAARPTFKOLITVFAERARPGRYLAI PGKFTRLPA-----Y 1169  
 QY 1006 RSLIEDD---DMGDLVDAEYLVPOQGFPCDPAPGAGMWHRRSSSTRSGGDLTLG 1062  
 DB 1170 TSGDEKILRKGLAFTDGSSEIAERPDYLOKRAAFPS-----HRTCT----- 1213  
 QY 1063 LBPSESEAP-----RSEPLASEGASDVFDG---DLGMAKAGLQSLPTHDPSPLOYS 1113  
 DB 1214 -----DELPKLNRYCKDPSNNKSSGTGDETDSSAREVGVNLR----- 1251  
 QY 1114 EDPVPLPSEFTDGYVAPLTCSPQPEYVNVQPDVRFQPPREGPLPAARPAATLERPKTL 1173  
 DB 1252 ----LDLPVDEDDYLMF--TCQPGPNNNNIN-----NPNONMAAVGAAGYM----- 1294  
 QY 1174 SPGNQGVYQVFAFGAAYENPEYL-----TPQGAAPOH-----PP 1211  
 DB 1295 -----DLIGVPVSYDNEBYLNAQTLGVSSPIPTQITGLFVWGVPMTMEYKVDMP 1345  
 QY 1212 AFSP-AFNDLYYMD 1224  
 DB 1346 GSEPTSSDHEYYND 1359

## RESULT 13

QY 09UK79 PRELIMINARY; PRT; 419 AA.  
 AC 09UK79;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Herstatin.  
 GN HER-2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94415951; PubMed=10485918;  
 RA Doherty J.K., Bond C., Jardim A., Adelman J.P., Clinton G.M.;  
 RT "The HER-2/neu receptor tyrosine kinase gene encodes a secreted  
 RT autoinhibitor";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:10869-10874 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Doherty J.K., Clinton G.M., Adelman J.P., Evans A.J., Hemmer W.D.;  
 RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF17761; AAD56009.2;  
 DR InterPro; IPR000494; EGFR\_L\_domain.  
 DR InterPro; IPR006211; Furin-Like.  
 DR InterPro; IPR006212; Furin-repeat.  
 DR Pfam; PF00757; Furin-like\_1.  
 DR Pfam; PF01030; Recep\_L\_domain\_1.  
 DR SMART; SMO0261; FU; 1.  
 SQ SEQUENCE 419 AA; 45472 MW; FECLBB347E2D030C CRC64;

Query Match 27.5%; Score 1874.5; DB 4; Length 419;  
 Best Local Similarity 89.6%; Pred. No. 8.9e-133;  
 Matches 353; Conservative 5; Mismatches 29; Indels 7; Gaps 2;

QY 1 MELALACMGLLALLPGCASTQVCTGDMKRLPASPEHLDMLRLHYGSCQVQVGNL 60  
 DB 1 MELALACMGLLALLPGCASTQVCTGDMKRLPASPEHLDMLRLHYGSCQVQVGNL 60  
 QY 61 ELTYLPTNASISFLQDIOEVQGYVLIANQVRQVPLQRLRIYRGTLQEDNVALAVLNG 120

DB 61 ELTYLPTNASISFLQDIOEVQGYVLIANQVRQVPLQRLRIYRGTLQEDNVALAVLNG 120  
 QY 121 DPLNNTPVTVGASPGGLRELOLRSLTEILKGVLIQRPOLCYODTILMKDIFKNNQLA 180  
 DB 121 DPLNNTPVTVGASPGGLRELOLRSLTEILKGVLIQRPOLCYODTILMKDIFKNNQLA 180  
 QY 181 LFLIDITNRSRACHQSPCKSRCKWSESSBECOSLTRVYCAAGGACRCKGPLETDCHEC 240  
 DB 181 LFLIDITNRSRACHQSPCKSRCKWSESSBECOSLTRVYCAAGGACRCKGPLETDCHEC 240  
 QY 241 AAGCTGPHSPCLALPHNSGICELHCPALVTYNTDFESMNPNEGRTFGASCYTACP 300  
 DB 241 AAGCTGPHSPCLALPHNSGICELHCPALVTYNTDFESMNPNEGRTFGASCYTACP 300  
 QY 301 YNLTSDVGSCTLYCPHNOEVTAEADGTORCEKSKPCARVYGLGMEHLREAVTSAN 360  
 DB 301 YNLTSDVGSCTLYCPHNOEVTAEADGTORCEKSKPCARVYGLGMEHLREAVTSAN 360  
 QY 361 IQEPAKCKLFGSLAFLEPSEFDGPASNTAPLQ 394  
 DB 356 LRQPG--PAHPVLSFLRPSWDLVSAFYSLEPLAP 387

## RESULT 14

QY 08R2X1 PRELIMINARY; PRT; 367 AA.  
 AC 08R2X1;  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Hypothetical 40.2 kDa protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Straubeberg R.;  
 RL Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; BC027080; AAH27080.1;  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR004019; YLP\_motif.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR Pfam; PF02757; YLP; 2.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 KW Hypothetical protein; ATP-binding; Transferase.  
 SQ SEQUENCE 367 AA; 40163 MW; 0BE03395F9E101B0 CRC64;

Query Match 25.6%; Score 1747; DB 11; Length 367;  
 Best Local Similarity 88.3%; Pred. No. 3e-123;  
 Matches 324; Conservative 11; Mismatches 32; Indels 0; Gaps 0;

QY 889 WMALESILRRRTTQSDVMSYGVYTWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPPIC 948  
 DB 1 WMALESILRRRTTQSDVMSYGVYTWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPPIC 948  
 QY 949 IDVYIMVWKWMDSECRPRELVESEFSRMAHDPQFVYIQNDLGG--PASPLDSTFY 1008  
 DB 61 IDVYIMVWKWMDSECRPRELVESEFSRMAHDPQFVYIQNDLGG--PASPLDSTFY 1008  
 QY 1009 LEDDDMDGLVDAEYLVPOQGFPCDPAPGAGMWHRRSSSTRSGGDLTLGLEPSEE 1068  
 DB 121 LEDDDMDGLVDAEYLVPOQGFPSDPAALGTSTAHRRHRSASAGGELTLGLEPSEE 180  
 QY 1069 EAPRSPLAPSSGAGSDVFDGDLGMAAGLQSLPHDPSPLQRYSEDPVPLPSEFTGYV 1128  
 DB 181 EPPRSPLAPSSGAGSDVFDGDLAVGTGLOSLSPHDLSPLOQRYSEDPVPLPSEFTGYV 240  
 QY 1129 APLTCSPPPEYVNOVDVQPPSPREGLPAARPAATLERPKTSLPGNGGVYQVFAFG 1188  
 DB 241 APLTCSPPPEYVNOVDVQPPSPREGLPAARPAATLERPKTSLPGNGGVYQVFAFG 300

```

QY 1189 GAVENPEYLTPGGAGAPQHPPEPPAFSPADNLYYWDOPPERGAPSTFKCTPAENPEY 1248
DB 301 GAVENPEYLAPRAGTASQPHSPAFSPADNLYYWDONSSQCGPPSTFECTPTAENPEY 360
QY 1249 LGLDVVP 1255
DB 361 LGLDVVP 367

RESULT 15
086712
ID 086712 PRELIMINARY; PRT; 729 AA.
AC 086712;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Polypeptide.
GN POLYPROTEIN.
OS Avian rous-associated virus type 1.
OC Viruses; Retroviridae; Retroviridae; Alpharetrovirus.
OX NCBI_Taxid=11950;
RN [1]
RP SEQUENCE FROM N. A.
RX MEDLINE=94203659; PubMed=8152791;
RA Vennetrom B., Raynosccheck C., Jansson L., Doederlein G., Lhotak V.,
RA Johansson A., Beug H.;
RT "Retroviral capture of c-erbB proto-oncogene sequences: rapid
RT evolution of distinct viral genomes carrying mutant v-erbB genes with
RT different transforming capacities.";
RL Oncogene 9:1307-1320(1994).
DR EMBL; S69372; AAC60725.1; -.
DR HSSP; P03322; 1A6S.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR004028; Retro_M.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; Dkinase; 1.
DR Pfam; PF02813; Retro_M_1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM_1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR_1.
KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 729 AA; 80649 MW; 84D2F6914EFID63 CRC64;

Query Match 25.2%; Score 1719; DB 15; Length 729;
Best Local Similarity 54.8%; Pred. No. 1,le-120;
Matches 358; Conservative 75; Mismatches 118; Indels 102; Gaps 15;
QY 569 PONGSVTCFGEADQVACAHYKDPFCVAPCPGSKPDLSTYMPIMKPPDEEGACOPCP 628
DB 141 PESTATRTKTPG--DHCKKCAHFIIDGPHCVKACPAAGVLENDTL--VMKADANANCOLCHP 197
QY 629 NCTHSCVDLDDKGCAPAGORASPLTISIAVY-GILLVVLGVVFGILKRQOKIRKXTM 687
DB 198 NCTRGCKAPGLEGP---NGSKTPSIAGVVGILLCLVVGILGLYLR--HIVRKTL 253
QY 688 RLLQETELVLPPLTPSGAMPQAOIKRIKTELKRVKVLGSGAGFTYKGIWIDGENYK 747
DB 254 RRLQERELVLPPLTPSGAPNQAHLRIKTEFKKVKVLGSGAGFTYKGIWIDGENYK 313
QY 748 IPVAIKYLRENTSPKANKELIDEAVYVMAVGSFYVSRLLGICLTSTVOLVQLMPYGL 807
DB 314 IPVAIKELRENTSPKANKELIDEAVYVMAVSVNPRVCRLLGICLTSTVOLVQLMPYGL 373
QY 808 DHVENRGRGLSGQDLLANWCQIAKMSYLEDEVRLVHRDLAARNVLVSPNHVKITDPELA 867
DB 374 DYIREHKONIGSYLILNWCQIAKGMNLYEERRLVHRDLAARNVLVKTPOHVKITDPELA 433
QY 868 RLLDIDETEVADGKVIKMALESLARRFTHOSDVWSYGTVVWELMTGAKPYDGI 927
DB 434 KLLGADEKRYHAEGKVPKIMWALESLIHLRIYTHOSDVWSYGTVVWELMTGSKPYDGI 493

```

```

QY 928 AREIPPLEKGERLPPPICTIDVYIMVCMWIDSECRPRELVSFARMADPORFV 987
DB 494 ASEISSVLEKGERLPPPICTIDVYIMVCMWIDSECRPRELVSFARMADPORFV 553
QY 988 VIO-NEDLPASPLDSTFRSLLEDMDGLVDAEYLVQOGFPCDPAPGAGMYHHR 1046
DB 554 VIOGERMHLPSPTDSKFYRTLMBERMDIVDAEYLVPHQGF 598
QY 1047 HRSSSTRSGGGLTLGLEPSEEAAPRPL-----APSEGASVDFDGLGMAKGLQSL 1101
DB 599 -NSPT-----SRTPLSSLSATSNMSTATNCID-----RNOQH 631
QY 1102 PTHDPSPLQRYSDPTVPLPSET--DGYAPLTCSPQPEYVNOVDVAPQPSREGLPA 1159
DB 632 PYVEDSFVQRYSDPTGNFLSESIDDQFL-----PAPEVNQ--LMPKKPS 675
QY 1160 APAGATLERPKTLSPGKGVVADV-----ATCGAVENPEYL 1197
DB 676 -----TAVQNOQIYNNISLTAISKLPMSRYQNSHSTAVDNEYL 715

```

Search completed: February 9, 2004, 16:16:45  
Job time : 52 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2004, 13:26:19 ; Search time 41 Seconds  
(without alignments)  
38.714 Million cell updates/sec

Title: US-09-930-125-3

Perfect score: 54

Sequence: 1 EBYLVPQGGF 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A-Geneseq\_19Jun03.\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match Length | ID              | Description         |
|------------|-------|--------------------|-----------------|---------------------|
| 1          | 54    | 100.0              | 10 23 AA20480   | Naturally processed |
| 2          | 54    | 100.0              | 10 23 AA077115  | Human Her-2/neu im  |
| 3          | 54    | 100.0              | 59 23 AAB21202  | Human HER-2/neu pr  |
| 4          | 54    | 100.0              | 58 23 AAM51147  | Human HER-2/neu pr  |
| 5          | 54    | 100.0              | 366 23 AAB21201 | Human HER-2/neu pr  |
| 6          | 54    | 100.0              | 165 23 AAM51146 | Human Her-2/neu on  |
| 7          | 54    | 100.0              | 283 23 AAB58875 | Breast and ovarian  |
| 8          | 54    | 100.0              | 583 23 AAE20483 | Human protein for   |
| 9          | 54    | 100.0              | 581 23 AA20481  | Human protein for   |

|    |    |       |                  |                    |
|----|----|-------|------------------|--------------------|
| 10 | 54 | 100.0 | 1589 23 AA20484  | Human protein for  |
| 11 | 54 | 100.0 | 600 23 AA20482   | Human protein for  |
| 12 | 54 | 100.0 | 712 21 AAB21204  | Human HER-2/neu fu |
| 13 | 54 | 100.0 | 712 21 AAM51149  | Her-2/neu extracel |
| 14 | 54 | 100.0 | 919 21 AAB21203  | Human HER-2/neu fu |
| 15 | 54 | 100.0 | 919 23 AAM51148  | Her-2/neu extracel |
| 16 | 54 | 100.0 | 920 23 AAM51152  | Mouse Her-2/neu ex |
| 17 | 54 | 100.0 | 926 23 AAM51153  | Mouse Her-2/neu ex |
| 18 | 54 | 100.0 | 1900 21 AAB21208 | Human HER-2/neu pr |
| 19 | 54 | 100.0 | 1823 23 AA098923 | Human breast cance |
| 20 | 54 | 100.0 | 1855 17 AAU01111 | HER-2/neu protein  |
| 21 | 54 | 100.0 | 1455 20 AAW92406 | Human HER-2/neu pr |
| 22 | 54 | 100.0 | 1155 21 AAB21198 | Human HER-2/neu pr |
| 23 | 54 | 100.0 | 1155 21 AA784780 | Amino acid sequenc |
| 24 | 54 | 100.0 | 1455 21 AA92620  | Human heregulin 2  |
| 25 | 54 | 100.0 | 1255 22 AAB21210 | Human tyrosine kin |
| 26 | 54 | 100.0 | 1255 22 AAB85458 | Human HER-2/neu pr |
| 27 | 54 | 100.0 | 1255 22 AAG88267 | HER2/neu amino aci |
| 28 | 54 | 100.0 | 1255 22 AAB60167 | HER2 transgene pla |
| 29 | 54 | 100.0 | 1255 23 AAE26349 | Human HER-2 protei |
| 30 | 54 | 100.0 | 1235 23 AAE26366 | Human HER2 antigen |
| 31 | 54 | 100.0 | 1235 23 AAB24067 | Human Her-2 protei |
| 32 | 54 | 100.0 | 1235 23 AA20479  | Human Her-2/neu pr |
| 33 | 54 | 100.0 | 1235 23 AAM51143 | Human Her-2/neu on |
| 34 | 54 | 100.0 | 1255 23 AAU77114 | Human Her-2/neu po |
| 35 | 54 | 100.0 | 1255 23 AAU74545 | Human HER2 (ErbB2) |
| 36 | 54 | 100.0 | 1255 24 ABR47447 | Breast cancer asso |
| 37 | 54 | 100.0 | 1255 24 ABP74708 | Rat HER-2/neu prot |
| 38 | 54 | 100.0 | 1255 21 AAB21199 | Mouse Her-2/neu pr |
| 39 | 54 | 100.0 | 1255 21 AAB21206 | Mouse Her-2/neu pr |
| 40 | 54 | 100.0 | 1255 22 AAG2860  | Amino acid sequenc |
| 41 | 54 | 100.0 | 1256 23 AAM51151 | Rat Her-2/neu onco |
| 42 | 54 | 100.0 | 1256 23 AAM51151 | Mouse Her-2/neu on |
| 43 | 54 | 100.0 | 1433 14 AAR39568 | Sequence of c-erbB |
| 44 | 54 | 92.6  | 11 14 AAR44876   | Inhibitor of TK of |
| 45 | 50 | 92.6  | 13 16 AAR73983   | EGF-R derived pept |

#### ALIGNMENTS

RESULT 1  
AA20480 standard; peptide; 10 AA.

AA20480; (first entry)

01-JUL-2002 (first entry)  
Naturally processed HLA-B\*44-restricted epitope of human Her-2/neu.  
Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;  
human leukocyte antigen; HLA; vaccine; malignancy; cytotoxic.

Homo sapiens.

MO200214503-A2

21-FEB-2002.

14-AUG-2001; 2001WO-US41733.

14-AUG-2000; 2000US-225152P.

28-SEP-2000; 2000US-236428P.

21-FEB-2001; 2001US-270520P.

(CORI-) CORIYA CORP.

Hand-zimmermann S, Cheever MA, Foy TM, Lodes MD, Kalos MD;  
McNeill PD, Veldick TS;  
WPI; 2002-280758/32.

PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,  
 PT prevention and diagnosis of cancer, preferably breast cancer -  
 PS Claim 2; Page 87; 129pp; English.  
 XX The invention relates to an isolated Her-2/Neu polypeptide composition  
 CC effective for eliciting an immune response. The invention is useful for  
 CC eliciting an immune response in a patient, where the patient is human  
 CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.  
 CC The composition is useful for the therapy and diagnosis of cancer.  
 CC Preferably breast cancer, in pharmaceutical compositions, e.g., vaccine  
 CC and other compositions for the diagnosis, prevention and treatment of  
 CC human malignancies, for stimulating and/or expanding T cells specific for  
 CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a  
 CC patient. The invention is useful for stimulating a T cell response in a  
 CC human patient, as probe or primer for nucleic acid hybridisation, to  
 CC selectively form duplex molecules with complementary stretches of the  
 CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full  
 CC length gene from a suitable library, and to direct expression of a  
 CC polypeptide in appropriate host cells. The composition is useful in  
 CC prophylactic or therapeutic applications and for the treatment of cancer,  
 CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-  
 CC associated malignancies. The invention is useful in gene therapy. The  
 CC present sequence is naturally processed HLA-B44-restricted epitope of  
 CC human Her-2/Neu protein.  
 SQ Sequence 10 AA:  
 Query Match 100.0%; Score 54; DB 23; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0041;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 EEYLVPOQGF 10  
 |||||  
 1 EEYLVPOQGF 10  
 Db 1 EEYLVPOQGF 10  
 RESULT 2  
 AAU77115  
 ID AAU77115 standard; Peptide: 10 AA.  
 AC AAU77115;  
 DT 05-JUN-2002 (first entry)  
 DE Human Her-2/neu immunogenic epitope.  
 KW Human; Her-2/neu; cytostatic; haematological malignancy; CML;  
 KM acute myelogenous leukaemia; AML; chronic myelogenous leukaemia; CLL;  
 KM chronic lymphocytic leukaemia; myeloma; non-Hodgkin's lymphoma; MDS;  
 KM Hodgkin's lymphoma; T cell therapy.  
 OS Homo sapiens.  
 PN WO200213847-A2.  
 PD 21-FEB-2002.  
 PF 13-AUG-2001; 2001WO-US25408.  
 PR 14-AUG-2000; 2000US-0638280.  
 PR 28-SEP-2000; 2000US-0675904.  
 PA (CORI-) CORIXA CORP.  
 PI Galger A, Cheever MA, Hand-zimmermann S;  
 DR WPI; 2002-280741/32.  
 PT Inhibiting haematological malignancy development by administering  
 PT polypeptide comprising immunogenic portion of Her-2/neu, polynucleotide  
 PT encoding the polypeptide, or antigen presenting cells expressing the  
 PT polypeptide

XX Claim 2; Page 49; 74pp; English.  
 XX The invention relates to a method for inhibiting development of  
 CC haematological malignancy in a patient by administering a polypeptide  
 CC comprising an immunogenic portion of Her-2/neu or a polynucleotide  
 CC encoding the polypeptide. Antigen presenting cells that express the  
 CC protein can also be administered. The sequences are used for inhibiting  
 CC development of haematological malignancy such as acute myelogenous  
 CC leukaemia (AML), chronic myelogenous leukaemia (CML), chronic lymphocytic  
 CC leukaemia (CLL), MDS, myelomas, Hodgkin's lymphoma and non-Hodgkin's  
 CC lymphoma. This sequence represents a human Her-2/neu immunogenic epitope.  
 SQ Sequence 10 AA:  
 Query Match 100.0%; Score 54; DB 23; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0041;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 EEYLVPOQGF 10  
 |||||  
 1 EEYLVPOQGF 10  
 Db 1 EEYLVPOQGF 10  
 RESULT 3  
 AAB21202  
 ID AAB21202 standard; protein: 59 AA.  
 AC AAB21202;  
 DT 12-JAN-2001 (first entry)  
 DE Human HER-2/neu protein phosphorylation domain partial sequence.  
 KW Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;  
 KM breast cancer; prostate cancer; ovarian cancer; lung cancer;  
 KM colon cancer.  
 OS Homo sapiens.  
 PN WO2000044899-A1.  
 PD 03-AUG-2000.  
 PF 28-JUN-2000; 2000WO-US02164.  
 PR 29-JAN-1999; 99US-0117976.  
 PA (CORI-) CORIXA CORP.  
 PA (SMIK) SMITHKLINE BEECHAM.  
 PI Cheever MA, Gheysen D;  
 DR WPI; 2000-505976/45.  
 PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins  
 PT useful for vaccinating against breast, ovarian, colon, lung and  
 PT prostate cancers -  
 PS Claim 27; Fig 11; 129pp; English.  
 XX The present sequence is a preferred portion of the phosphorylation domain  
 CC of the HER-2/neu protein. HER-2/neu is a member of the tyrosine kinase  
 CC family of receptor-like glycoproteins and shows homology to the epidermal  
 CC growth factor receptor (EGFR). It probably plays a part in cell growth  
 CC and/or differentiation. The HER-2/neu gene is an oncogene. An HER-2/neu  
 CC fusion protein comprising a HER-2/neu extracellular domain fused to a  
 CC HER-2/neu phosphorylation domain may be used to treat or prevent cancer  
 CC by eliciting or enhancing an immune response to the HER-2/neu protein. It  
 CC may be used to treat malignancies such as breast, ovarian, colon, lung  
 CC and prostate cancers, and may be used as an antigen to vaccinate against  
 CC these neoplasias.

*Handwritten signature: P6 X DP*

SQ Sequence 59 AA;  
 Query Match 100.0%; Score 54; DB 21; Length 59;  
 Best Local Similarity 100.0%; Pred. No. 0.023;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEYLVPOQGF 10  
 |||||  
 DB 32 EEYLVPOQGF 41  
 RESULT 4  
 ID AAM51147 standard; Protein; 59 AA.  
 XX AAM51147;  
 AC AAM51147;  
 XX 17-JUN-2002 (first entry)  
 DT 17-JUN-2002 (first entry)  
 DE Human Her-2/neu oncoprotein phosphorylation domain fragment.  
 XX Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;  
 KM tyrosine kinase; receptor; c-erbB2; gene therapy.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 PN WO200212341-A2.  
 PD 14-FEB-2002.  
 XX 03-AUG-2001; 2001WO-US24283.  
 PF 03-AUG-2001; 2000US-0632507.  
 XX 03-AUG-2000; 2000US-0632507.  
 PR 03-AUG-2000; 2000US-0632507.  
 XX (CORI-) CORIXA CORP.  
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX Cheever MA, Gheyssen D;  
 PI WPI; 2002-241743/29.  
 DR WPI; 2002-241743/29.  
 XX Her-2/neu fusion protein for treating or preventing cancer by eliciting  
 PT or enhancing an immune response to the protein, has Her-2/neu  
 PT extracellular domain fused to Her-2/neu intracellular or  
 PT phosphorylation domain -  
 XX Disclosure; Fig 11; 14pp; English.  
 PS The present sequence is that of a fragment (DeltapD) of the  
 CC phosphorylation domain of human Her-2/neu (p185 glycoprotein,  
 CC c-erbB2), an oncogenic self-protein and target for anti-cancer  
 CC vaccines. The Her-2/neu gene is amplified and p185 is  
 CC overexpressed in breast, ovarian, colon, lung and prostate cancer.  
 CC Her-2/neu (see AAM51143) is a member of the tyrosine kinase family  
 CC of receptor-like glycoproteins. It comprises an extracellular  
 CC domain with homology to the epidermal growth factor receptor  
 CC (EGFR), a highly hydrophobic transmembrane domain and a C-terminal  
 CC intracellular domain that also shows homology to EGFR. Its  
 CC overexpression correlates with a poor prognosis in breast and  
 CC ovarian cancers. The invention provides Her-2/neu fusion  
 CC proteins, nucleic acids encoding them, viral vectors, and vaccines  
 CC comprising the fusion proteins or nucleic acid molecules. In  
 CC preferred fusion proteins, the extracellular domain of a Her-2/neu  
 CC protein is fused to a Her-2/neu intracellular domain or  
 CC phosphorylation domain (or the DeltapD fragment). An immune  
 CC response to Her-2/neu protein is elicited or enhanced by  
 CC administering the fusion protein in the form of a vaccine, or by  
 CC transfecting cells of an animal *ex vivo* with a nucleic acid  
 CC encoding the fusion protein, and delivering the transfected cells  
 CC to the animal. The fusion proteins, nucleic acids, and isolated  
 CC specific T-cells are useful for inhibiting the development of a  
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer  
 CC in a patient. T cells that specifically react with a Her-2/neu

CC fusion protein can be used to remove tumour cells from a sample in  
 CC order to inhibit the development of cancer in a patient.  
 XX  
 SQ Sequence 59 AA;  
 Query Match 100.0%; Score 54; DB 23; Length 59;  
 Best Local Similarity 100.0%; Pred. No. 0.023;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEYLVPOQGF 10  
 |||||  
 DB 32 EEYLVPOQGF 41  
 RESULT 5  
 ID AAB21201 standard; protein; 266 AA.  
 XX AAB21201  
 AC AAB21201;  
 XX 12-JAN-2001 (first entry)  
 DT 12-JAN-2001 (first entry)  
 DE Human HER-2/neu protein phosphorylation domain.  
 XX Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;  
 KM breast cancer; prostate cancer; ovarian cancer; lung cancer;  
 KM colon cancer.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 PN WO200044899-A1.  
 PD 03-AUG-2000.  
 XX 28-JAN-2000; 2000WO-US02164.  
 PF 28-JAN-2000; 2000WO-US02164.  
 XX 29-JAN-1999; 99US-0117976.  
 PR 29-JAN-1999; 99US-0117976.  
 XX (CORI-) CORIXA CORP.  
 PA (SMIK) SMITHKLINE BEECHAM.  
 XX Cheever MA, Gheyssen D;  
 PI WPI; 2000-505976/45.  
 DR WPI; 2000-505976/45.  
 XX HER-2/neu extracellular domain/phosphorylation domain fusion proteins  
 PT useful for vaccinating against breast, ovarian, colon, lung and  
 PT prostate cancers -  
 XX Claim 2; Fig 10; 128pp; English.  
 PS The present sequence is the phosphorylation domain of the HER-2/neu  
 CC protein. HER-2/neu is a member of the tyrosine kinase family of  
 CC receptor-like glycoproteins and shows homology to the epidermal growth  
 CC factor receptor (EGFR). It probably plays a part in cell growth and/or  
 CC differentiation. The HER-2/neu gene is an oncogene. An HER-2/neu fusion  
 CC protein comprising a HER-2/neu extracellular domain fused to a HER-2/neu  
 CC phosphorylation domain may be used to treat or prevent cancer by  
 CC eliciting or enhancing an immune response to the HER-2/neu protein. It  
 CC may be used to treat malignancies such as breast, ovarian, colon, lung  
 CC and prostate cancers, and may be used as an antigen to vaccinate against  
 CC these neoplasias.  
 XX  
 SQ Sequence 266 AA;  
 Query Match 100.0%; Score 54; DB 21; Length 266;  
 Best Local Similarity 100.0%; Pred. No. 0.1;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEYLVPOQGF 10  
 |||||  
 DB 32 EEYLVPOQGF 41



RESULT 6  
ID AAM51146 standard; Protein; 266 AA.  
XX  
AC AAM51146;  
XX  
DT 17-JUN-2002 (first entry)  
XX  
DE Human Her-2/neu oncoprotein phosphorylation domain.  
XX  
KW Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;  
KW tyrosine kinase; receptor; c-erbB2; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN MO200212341-A2.  
XX  
PD 14-FEB-2002.  
XX  
PF 03-AUG-2001; 2001MO-US24283.  
XX  
PR 03-AUG-2000; 2000US-0632507.  
XX  
PA (CORI-) CORIAX CORP.  
PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.  
PI Cheever MA, Cheysen D;  
DR WPI; 2002-241743/29.  
XX  
PT Her-2/neu fusion protein for treating or preventing cancer by eliciting  
PT or enhancing an immune response to the protein, has Her-2/neu  
PT extracellular domain fused to Her-2/neu intracellular or  
PT phosphorylation domain -  
XX  
PS Claim 2; Fig 10; 141pp; English.  
XX  
CC The present sequence is that of the phosphorylation domain of  
CC human Her-2/neu (p185 glycoprotein or c-erbB2), an oncogenic  
CC self-protein and target for anti-cancer vaccines. The Her-2/neu  
CC gene is amplified and p185 is overexpressed in a variety of cancers,  
CC including breast, ovarian, colon, lung and prostate cancer.  
CC Her-2/neu (see AAM51143) is a member of the tyrosine kinase family  
CC of receptor-like glycoproteins. It comprises an extracellular  
CC domain with homology to the epidermal growth factor receptor  
CC (EGFR), a highly hydrophobic transmembrane domain and a C-terminal  
CC intracellular domain that also shows homology to EGFR. Its  
CC overexpression correlates with a poor prognosis in breast and  
CC ovarian cancers. The invention provides Her-2/neu fusion  
CC proteins, nucleic acids encoding them, viral vectors, and vaccines  
CC comprising the fusion proteins or nucleic acid molecules. In  
CC preferred fusion proteins, the extracellular domain of a Her-2/neu  
CC protein is fused to a Her-2/neu intracellular domain or  
CC phosphorylation domain (or its DeltapD fragment). An immune  
CC response to Her-2/neu protein is elicited or enhanced by  
CC administering the fusion protein in the form of a vaccine, or by  
CC transfecting cells of an animal *ex vivo* with a nucleic acid  
CC encoding the fusion protein, and delivering the transfected cells  
CC to the animal. The fusion proteins, nucleic acids, and isolated  
CC specific T-cells are useful for inhibiting the development of a  
CC cancer, especially breast, ovarian, colon, lung or prostate cancer  
CC in a patient. T cells that specifically react with a Her-2/neu  
CC fusion protein can be used to remove tumour cells from a sample in  
CC order to inhibit the development of cancer in a patient.  
XX  
SQ Sequence 266 AA;

Query Match 100.0%; Score 54; DB 23; Length 266;  
Best Local Similarity 100.0%; Pred. No. 0.1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 EBYLVPQGGF 10

Db 32 EBYLVPQGGF 41  
RESULT 7  
ID AAB58875 standard; Protein; 293 AA.  
XX  
AC AAB58875;  
XX  
DT 27-MAR-2001 (first entry)  
XX  
DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 583.  
XX  
KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;  
KW nocotropic; neuroprotective; antiviral; antiallergic; hepatotropic;  
KW antidiabetic; antiinflammatory; antiviral; vulnery; anticonvulsant;  
KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;  
KW Addison's disease; allergy; autoimmune haemolytic anaemia;  
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;  
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;  
KW cardiovascular disorder; wound healing; neurological disease.  
XX  
OS Homo sapiens.  
XX  
PN MO200055173-A1.  
XX  
PD 21-SEP-2000.  
XX  
PE 08-MAR-2000; 2000MO-US05881.  
XX  
PR 12-MAR-1999; 99US-0124270.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM;  
XX  
DR WPI; 2000-611515/58.  
XX  
DT N-PSDB; AAF21778.  
XX

PT New human breast and ovarian cancer associated gene sequences and the  
PT polypeptides encoded by these genes, useful in the prevention,  
PT treatment and diagnosis of cancer, immune disorders, cardiovascular  
PT disorders and neurological diseases -  
XX  
PS Claim 11; Page 1019-1020; 1299pp; English.

CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are  
CC associated with breast and ovarian cancer. Included in the invention are  
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
CC isolation and characterisation of the DNA and protein sequences of the  
CC invention. The breast and ovarian cancer associated DNA, protein, agonist  
CC or antagonist sequences exhibit cytostatic; immunosuppressive;  
CC nocotropic; neuroprotective; antiviral; antiallergic; hepatotropic;  
CC antidiabetic; antiinflammatory; antiviral; vulnery; anticonvulsant;  
CC antibacterial; antifungal; antiparasitic and cardiac activity. The  
CC polynucleotide and protein sequences are used in the diagnosis of cancer,  
CC particularly breast and ovarian cancer. The nucleic acid sequences,  
CC proteins, agonists and antagonists may also be used in the diagnosis,  
CC prevention and treatment of immune disorders e.g. Addison's disease,  
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
CC arthritis and ulcerative colitis; cardiovascular disorders such as  
CC myocardial ischaemia; wound healing; neurological diseases such as  
CC cerebral ischaemia and epilepsy; and infectious diseases.  
XX  
SQ Sequence 293 AA;

Query Match 100.0%; Score 54; DB 21; Length 293;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```
QY      1 EBYLVPOGGF 10
         |||||
Db      105 EBYLVPOGGF 114
```

|          |                                    |
|----------|------------------------------------|
| RESULT 8 |                                    |
| AAE20483 |                                    |
| ID       | AAE20483 standard; Protein; 583 AA |

| QY | 1   | BEYLPQGGF | 10  |
|----|-----|-----------|-----|
| Db | 347 | BEYLPQGGF | 356 |

| RESULT 9                              |
|---------------------------------------|
| AAE20481                              |
| ID AAE20481 standard; Protein; 587 AA |

Db 347 EEVLVPQGGF 356

RESULT 10  
ID AAE20484 standard; Protein; 589 AA.  
AC AAE20484;  
DT 01-JUL-2002 (first entry)  
DE Human protein for clone HICD\_in\_gpdm\_coding\_sequence.  
KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer; human leukocyte antigen; HLA; vaccine; malignancy; cytostatic.  
OS Homo sapiens.  
PN WO200214503-A2.  
PD 21-FEB-2002.  
PF 14-AUG-2001; 2001WO-US41733.  
PR 14-AUG-2000; 2000US-225152P.  
PR 28-SEP-2000; 2000US-236428P.  
PR 21-FEB-2001; 2001US-270520P.  
PA (CORI-) CORIXA CORP.  
PI Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;  
PI McNeill PD, Vedvick TS;  
DR WPI; 2002-280758/32.  
DR N-PSDB; AAD32747.  
PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,  
PT prevention and diagnosis of cancer, preferably breast cancer -  
PS Example 5; Page 122-123; 129pp; English.  
XX The invention relates to an isolated Her-2/Neu polypeptide composition  
XX effective for eliciting an immune response. The invention is useful for  
XX eliciting an immune response in a patient, where the patient is human  
XX leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.  
XX The composition is useful for the therapy and diagnosis of cancer,  
XX preferably breast cancer, in pharmaceutical compositions, e.g., vaccine  
XX and other compositions for the diagnosis, prevention and treatment of  
XX human malignancies, for stimulating and/or expanding T cells specific for  
XX Her-2/Neu polypeptide and for inhibiting the development of cancer in a  
XX patient. The invention is useful for stimulating a T cell response in a  
XX human patient, as probe or primer for nucleic acid hybridization, to  
XX selectively form duplex molecules with complementary stretches of the  
XX entire Her-2/Neu gene or gene fragments of interest, to isolate a full  
XX length gene from a suitable library, and to direct expression of a  
XX polypeptide in appropriate host cells. The composition is useful in  
XX prophylactic or therapeutic applications and for the treatment of cancer,  
XX preferably for the immunotherapy of breast cancer and other Her-2/Neu-  
XX associated malignancies. The invention is useful in gene therapy. The  
XX present sequence is human protein for the clone HICD\_in\_gpdm\_coding\_  
XX sequence.  
SQ Sequence 589 AA;

Query Match 100.0%; Score 54; DB 23; Length 589;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVLVPQGGF 10  
DB 355 EEVLVPQGGF 364

RESULT 11

ID AAE20482 standard; Protein; 600 AA.  
AC AAE20482;  
DT 01-JUL-2002 (first entry)  
DE Human protein for the clone HICD\_plus\_8\_HIS.  
KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer; human leukocyte antigen; HLA; vaccine; malignancy; cytostatic.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Misc-difference 515  
FT /note="Encoded by GCC"  
PN WO200214503-A2.  
PD 21-FEB-2002.  
PF 14-AUG-2001; 2001WO-US41733.  
PR 14-AUG-2000; 2000US-225152P.  
PR 28-SEP-2000; 2000US-236428P.  
PR 21-FEB-2001; 2001US-270520P.  
PA (CORI-) CORIXA CORP.  
PI Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;  
PI McNeill PD, Vedvick TS;  
DR WPI; 2002-280758/32.  
DR N-PSDB; AAD32745.  
PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,  
PT prevention and diagnosis of cancer, preferably breast cancer -  
PS Example 5; Page 124-125; 129pp; English.  
XX The invention relates to an isolated Her-2/Neu polypeptide composition  
XX effective for eliciting an immune response. The invention is useful for  
XX eliciting an immune response in a patient, where the patient is human  
XX leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.  
XX The composition is useful for the therapy and diagnosis of cancer,  
XX preferably breast cancer, in pharmaceutical compositions, e.g., vaccine  
XX and other compositions for the diagnosis, prevention and treatment of  
XX human malignancies, for stimulating and/or expanding T cells specific for  
XX Her-2/Neu polypeptide and for inhibiting the development of cancer in a  
XX patient. The invention is useful for stimulating a T cell response in a  
XX human patient, as probe or primer for nucleic acid hybridization, to  
XX selectively form duplex molecules with complementary stretches of the  
XX entire Her-2/Neu gene or gene fragments of interest, to isolate a full  
XX length gene from a suitable library, and to direct expression of a  
XX polypeptide in appropriate host cells. The composition is useful in  
XX prophylactic or therapeutic applications and for the treatment of cancer,  
XX preferably for the immunotherapy of breast cancer and other Her-2/Neu-  
XX associated malignancies. The invention is useful in gene therapy. The  
XX present sequence is human protein for the clone HICD\_plus\_8\_HIS.  
SQ Sequence 600 AA;

Query Match 100.0%; Score 54; DB 23; Length 600;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVLVPQGGF 10  
DB 366 EEVLVPQGGF 375

RESULT 12  
 AAB21204  
 ID AAB21204 standard; protein; 712 AA.  
 AC AAB21204;  
 XX  
 DT 12-JAN-2001 (first entry)  
 XX  
 DE Human HER-2/neu fusion protein.  
 XX  
 KW Human; HER-2/neu; oncogene; tyrosine kinase; cytosolic; vaccine;  
 KW breast cancer; prostate cancer; ovarian cancer; lung cancer;  
 KW colon cancer; fusion protein.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200044899-A1.  
 XX  
 PD 03-AUG-2000.  
 XX  
 PF 28-JAN-2000; 2000MO-US02164.  
 XX  
 PR 29-JAN-1999; 99US-0117976.  
 XX  
 PA (CORI-) CORIXA CORP.  
 PA (SMIK) SMITHKLINE BEECHAM.  
 XX  
 PI Cheever MA, Gheyssen D;  
 XX  
 DR WPI; 2000-505976/45.  
 XX  
 PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins  
 PT useful for vaccinating against breast, ovarian, colon, lung and  
 PT prostate cancers -  
 XX  
 PS Claim 27; Fig 13; 128pp; English.  
 XX  
 CC The present sequence is a fusion protein comprising the extracellular  
 CC domain and a preferred portion of the phosphorylation domain of the human  
 CC HER-2/neu protein. HER-2/neu is a member of the tyrosine kinase family of  
 CC receptor-like glycoproteins and shows homology to the epidermal growth  
 CC factor receptor (EGFR). It probably plays a part in cell growth and/or  
 CC differentiation. The HER-2/neu gene is an oncogene. HER-2/neu fusion  
 CC proteins may be used to treat or prevent cancer by eliciting or enhancing  
 CC an immune response to the HER-2/neu protein. They may be used to treat  
 CC malignancies such as breast, ovarian, colon, lung and prostate cancers,  
 CC and may be used as an antigen to vaccinate against these neoplasias.  
 CC  
 XX  
 SQ Sequence 712 AA;  
 Query Match 100.0%; Score 54; DB 21; Length 712;  
 Best Local Similarity 100.0%; Pred. No. 0.26;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EBYLVPOQGF 10  
 DB 685 EBYLVPOQGF 694  
 RESULT 13  
 AAM51149  
 ID AAM51149 standard; protein; 712 AA.  
 AC AAM51149;  
 XX  
 DT 17-JUN-2002 (first entry)  
 XX  
 DE Her-2/neu extracellular domain-delta-phosphorylation domain fusion.  
 XX  
 KW Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;  
 KW tyrosine kinase; receptor; c-erbB2; gene therapy.  
 XX

OS Homo sapiens.  
 XX  
 FH Key  
 FT Domain  
 FT 1..653  
 FT /note= "extracellular domain"  
 FT Domain  
 FT 654..712  
 FT /note= "phosphorylation domain fragment"  
 XX  
 PN WO200212341-A2.  
 XX  
 PD 14-FEB-2002.  
 XX  
 PF 03-AUG-2001; 2001WO-US24283.  
 XX  
 PR 03-AUG-2000; 2000US-0632507.  
 XX  
 PA (CORI-) CORIXA CORP.  
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Cheever MA, Gheyssen D;  
 XX  
 DR WPI; 2002-241743/29.  
 XX  
 PT Her-2/neu fusion protein for treating or preventing cancer by eliciting  
 PT or enhancing an immune response to the protein, has Her-2/neu  
 PT extracellular domain fused to Her-2/neu intracellular or  
 PT phosphorylation domain -  
 XX  
 PS Claim 37; Fig 13; 14pp; English.  
 XX  
 CC The present sequence is that of a fusion protein between the  
 CC extracellular domain and a fragment (DeltaEPD) of the phosphorylation  
 CC domain of human Her-2/neu (see AAM51143), an oncogenic self-protein  
 CC and target for anti-cancer vaccines. The fusion protein can be  
 CC obtained by recombinant DNA methods. Her-2/neu overexpression  
 CC correlates with a poor prognosis in breast and ovarian cancers.  
 CC The invention provides Her-2/neu fusion proteins, nucleic acids  
 CC encoding them, viral vectors, and vaccines comprising the fusion  
 CC proteins or nucleic acid molecules. In preferred fusion proteins,  
 CC the extracellular domain of Her-2/neu is fused to a Her-2/neu  
 CC intracellular domain or phosphorylation domain (or its DeltaEPD  
 CC fragment). An immune response to Her-2/neu protein is elicited or  
 CC enhanced by administering the fusion protein in the form of a vaccine,  
 CC or by transfecting cells of an animal ex vivo with a nucleic acid  
 CC encoding the fusion protein, and delivering the transfected cells  
 CC to the animal. The fusion proteins, nucleic acids, and isolated  
 CC specific T-cells are useful for inhibiting the development of a  
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer  
 CC in a patient. T cells that specifically react with a Her-2/neu  
 CC fusion protein can be used to remove tumour cells from a sample in  
 CC order to inhibit the development of cancer in a patient.  
 CC  
 XX  
 SQ Sequence 712 AA;  
 Query Match 100.0%; Score 54; DB 23; Length 712;  
 Best Local Similarity 100.0%; Pred. No. 0.26;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EBYLVPOQGF 10  
 DB 685 EBYLVPOQGF 694  
 RESULT 14  
 AAB21203  
 ID AAB21203 standard; protein; 919 AA.  
 AC AAB21203;  
 XX  
 DT 12-JAN-2001 (first entry)  
 XX  
 DE Human HER-2/neu fusion protein.  
 XX

KW Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;  
 KW breast cancer; prostate cancer; ovarian cancer; lung cancer;  
 KW colon cancer; fusion protein.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN W0200044899-A1.  
 XX  
 PD 03-AUG-2000.  
 XX  
 PF 28-JAN-2000; 2000WO-US02164.  
 XX  
 PR 29-JAN-1999; 99US-0117976.  
 XX  
 PA (CORI-) CORIXA CORP.  
 PA (SMIK ) SMITHKLINE BEECHAM.  
 XX  
 PI Cheever MA, Gheysen D;  
 XX  
 DR WPI; 2000-505976/45.  
 XX  
 PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins  
 PT useful for vaccinating against breast, ovarian, colon, lung and  
 PT prostate cancers -  
 XX  
 PS Claim 2; Fig 12; 128pp; English.  
 XX  
 CC The present sequence is a fusion protein comprising the extracellular  
 CC domain and the phosphorylation domain of the human HER-2/neu protein.  
 CC HER-2/neu is a member of the tyrosine kinase family of receptor-like  
 CC glycoproteins and shows homology to the epidermal growth factor receptor  
 CC (EGFR). It probably plays a part in cell growth and/or differentiation.  
 CC The HER-2/neu gene is an oncogene. HER-2/neu fusion proteins may be used  
 CC to treat or prevent cancer by eliciting or enhancing an immune response  
 CC to the HER-2/neu protein. They may be used to treat malignancies such as  
 CC breast, ovarian, colon, lung and prostate cancers, and may be used as an  
 CC antigen to vaccinate against these neoplasias.  
 XX  
 SQ Sequence 919 AA;  
 XX  
 Query Match 100.0%; Score 54; DB 21; Length 919;  
 Best Local Similarity 100.0%; Pred. No. 0.34;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EBYLVPOQGF 10  
 DB 685 EBYLVPOQGF 694  
 XX  
 RESULT 15  
 ID AAM51148 standard; Protein; 919 AA.  
 XX  
 AC AAM51148;  
 XX  
 DT 17-JUN-2002 (first entry)  
 XX  
 DE Her-2/neu extracellular domain-phosphorylation domain fusion.  
 XX  
 KW Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;  
 KW tyrosine kinase; receptor; c-erbB2; gene therapy.  
 XX  
 OS Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 1..653  
 FT /note= "extracellular domain"  
 FT Domain 654..919  
 FT /note= "phosphorylation domain"  
 XX  
 PN W0200212341-A2.  
 XX

PD 14-FEB-2002.  
 XX  
 PF 03-AUG-2001; 2001WO-US24283.  
 XX  
 PR 03-AUG-2000; 2000US-0632507.  
 XX  
 PA (CORI-) CORIXA CORP.  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Cheever MA, Gheysen D;  
 XX  
 DR WPI; 2002-241743/29.  
 XX  
 PT Her-2/neu fusion protein for treating or preventing cancer by eliciting  
 PT or enhancing an immune response to the protein, has Her-2/neu  
 PT extracellular domain fused to Her-2/neu intracellular or  
 PT phosphorylation domain -  
 XX  
 PS Claim 2; Fig 12; 141pp; English.  
 XX  
 CC The present sequence is that of a fusion protein between the  
 CC extracellular domain and phosphorylation domain of human Her-2/neu  
 CC (see AAM51143), an oncogenic self-protein and target for anti-cancer  
 CC vaccines. The fusion protein can be obtained by recombinant DNA  
 CC methods. Her-2/neu overexpression correlates with a poor prognosis  
 CC in breast and ovarian cancers. The invention provides Her-2/neu  
 CC fusion proteins, nucleic acids encoding them, viral vectors, and  
 CC vaccines comprising the fusion proteins or nucleic acid molecules.  
 CC In preferred fusion proteins, the extracellular domain of a  
 CC Her-2/neu protein is fused to a Her-2/neu intracellular domain or  
 CC phosphorylation domain (or its DeltaCD fragment). An immune  
 CC response to Her-2/neu protein is elicited or enhanced by  
 CC administering the fusion protein in the form of a vaccine, or by  
 CC transfecting cells of an animal *ex vivo* with a nucleic acid  
 CC encoding the fusion protein, and delivering the transfected cells  
 CC to the animal. The fusion proteins, nucleic acids, and isolated  
 CC specific T-cells are useful for inhibiting the development of a  
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer  
 CC in a patient. T cells that specifically react with a Her-2/neu  
 CC fusion protein can be used to remove tumour cells from a sample in  
 CC order to inhibit the development of cancer in a patient.  
 XX  
 SQ Sequence 919 AA;  
 XX  
 Query Match 100.0%; Score 54; DB 23; Length 919;  
 Best Local Similarity 100.0%; Pred. No. 0.34;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EBYLVPOQGF 10  
 DB 685 EBYLVPOQGF 694  
 XX  
 Search completed: February 7, 2004, 13:31:17  
 Job time : 42 secs